



Machine Learning-Based Prediction of Firmness in Coated Bananas Under Retail Conditions

Mawande Shinga^{1,2} · Yardjouma Silue^{1,2} · Olaniyi Fawole^{1,2}

Received: 20 April 2025 / Accepted: 8 July 2025 / Published online: 12 July 2025
© The Author(s) 2025

Abstract

This study explored the use of machine learning (ML) algorithms for the non-destructive prediction of banana firmness under retail conditions, facilitating real-time quality assessment, optimising supply chain decisions, and improving postharvest management. Bananas were coated with the optimal formulation of *Opuntia ficus-indica* mucilage (OF) and stored at 23 ± 2 °C for 10 days. A factorial experimental design was employed, with edible coating and storage duration as primary factors. Banana parameters, including respiration rate, ethylene production, weight loss, and colour, were measured alongside firmness. Collected data was used to develop predictive models using ML techniques, namely Partial Least Squares (PLS) regression, Ridge regression, and Elastic Net regression. The results showed that banana firmness could be predicted using non-invasive attributes, with respiration rate and weight loss being the most influential predictors. Among the models tested, PLS regression exhibited the highest predictive accuracy, with an R^2 of 0.978, RMSE of 0.097, MAE of 0.009, and R^2 -adjusted value of 0.940. Ridge regression followed closely (R^2 of 0.972, RMSE of 0.110, MAE of 0.012, and R^2 -adjusted of 0.922), while Elastic Net regression, though slightly less precise, still demonstrated strong predictive capability ($R^2 = 0.956$, RMSE=0.142, MSE=0.020, R^2 -adjusted=0.801). This study also demonstrated that the application of an optimised *Opuntia ficus-indica* mucilage extended the shelf-life of bananas by four days. This approach allows real-time quality assessment, enhancing quality control, reducing postharvest losses, and improving inventory management in the fruit industry.

Keywords Machine Learning Algorithms · Non-invasive Parameters · Edible Coatings · Predictive Power · Firmness

Introduction

Fruit softening is an irreversible process caused by cell wall alterations due to the depolymerisation of cell wall polysaccharides [1–3]. These structural changes, particularly in cellulose organization and cell turgor, are closely associated with firmness, a critical indicator of fruit quality,

including bananas [4–6]. The firmness of bananas plays an important role in consumer acceptance, shelf-life, and overall marketability [7, 8]. This is because firmness is associated with freshness, ripeness, and eating quality. However, the progressive reduction in firmness accelerates ripening, increases susceptibility to pathogen invasion, and leads to significant postharvest losses [9, 10]. Changes in firmness during ripening are often accompanied by alterations in chemical composition, which in turn affect sensory attributes, nutritional quality, and consumer appeal [11–14]. To mitigate firmness loss in bananas, various postharvest treatments have been explored, including gaseous applications, physical treatments, and chemical agents [15–20].

Among postharvest technologies, edible coatings have gained increasing attention due to their biodegradability, ability to form semi-permeable barriers, and capacity to delay ripening by modulating gas exchange [14, 21–23]. Previous studies, including our own, have demonstrated that *Opuntia ficus-indica* mucilage (OFIM) can effectively

✉ Olaniyi Fawole
olaniyif@uj.ac.za

¹ Postharvest and Agroprocessing Research Centre,
Department of Botany and Plant Biotechnology,
University of Johannesburg, P.O. Box 524, Auckland Park,
Johannesburg 2006, South Africa

² South African Research Chairs Initiative in Sustainable
Preservation and Agroprocessing Research, Faculty of
Science, University of Johannesburg, P.O. Box 524, P.O. Box
524,
Auckland Park, Johannesburg, Auckland Park, Johannesburg
2006, South Africa

reduce banana firmness loss and extend shelf life [24]. OFIM also demonstrated its ability to form film, biodegradability, and effectiveness in modulating postharvest physiology. However, the effectiveness of edible coatings depends strongly on their composition and physicochemical properties, which must be carefully optimised to improve their functional performance. To address this, we conducted a follow-up study to optimize the OFIM coating by adjusting the concentrations of pomegranate peel extract (PPE), cellulose nanofiber (CNF), and glycerol. Each component plays a distinct role in enhancing the effectiveness of the coating in preserving banana firmness [25].

Given the importance of firmness in determining banana quality, traditional methods such as penetrometer and puncture tests are widely used for texture assessment [26]. However, these methods are inherently destructive, rendering the tested fruit unmarketable and contributing to food waste [27]. Consequently, the demand for non-invasive approaches has grown, particularly for large-scale quality control and supply chain optimisation [26, 28]. Several non-destructive techniques, including ultrasonic, acoustic vibration, optical imaging, and electrical capacitance methods, have been explored for fruit firmness measurement [29, 30]. However, these techniques are often limited by their accessibility, cost, and inability to predict firmness evolution over time, thereby restricting their practical application in dynamic postharvest environments. As a result, fruit handlers frequently rely on manual inspection and subjective sorting, which are labour-intensive and inconsistent for large-scale commercial operations.

The emergence of artificial intelligence (AI) technologies, particularly machine learning (ML) algorithms, presents a promising alternative for advancing postharvest management strategies. Machine learning, a subset of AI, focuses on developing computational models that learn from data and make predictions without explicit programming [31]. Various ML models, including artificial neural networks (ANN) [32], random forest (RF) [33], partial least squares (PLS) regression [28], support vector machines (SVM) [34], and Ridge regression [35], have been successfully applied in postharvest quality assessment. By integrating non-destructive measurements with machine learning algorithms makes it possible to predict firmness dynamically, enabling data-driven decision-making and proactive intervention strategies. This approach enhances supply chain efficiency by optimising sorting, inventory management, and quality control, ultimately reducing postharvest losses.

This study addresses the gap in predicting the firmness of edible-coated bananas using non-destructive attributes. It investigates the impact of an optimised OFIM-based edible coating on the banana quality and compares multiple ML

models, enhanced by hyperparameter tuning, to identify the most accurate approach for firmness prediction under retail conditions. The findings aim to support scalable, non-destructive postharvest monitoring strategies.

Materials and Methods

Fruit Material

Physiologically mature green banana (*Musa* spp) fruits (273 fingers) were transported from Umbhala Estates farm, Mpumalanga (25°25'48.7" S; 31°57'00.5" E), to the Postharvest and Agroprocessing Research Centre (26°11'3.9" S, 27°59'45.9" E), Johannesburg, South Africa. The fruit were visually inspected and sorted for defects, colour and size consistency. Subsequently, the bananas were disinfected with 0.1% sodium hypochlorite to eliminate potential contaminants.

Mucilage from *Opuntia Ficus Cladodes* Extraction

Cladodes sourced from the prickly pear orchard at Ubali Farm in Pretoria, Gauteng (25°38'31.9" S, 28°27'33.0" E). The mucilage was derived from mature cladodes of *Opuntia ficus-indica* according to the method described by Shinga and Fawole [24]. Briefly, the cladodes were peeled, cubed, and microwaved at 800 W for 4 min. Thereafter, the insoluble fibres were removed using cheesecloth, and the remaining liquid was centrifuged at 8000 x g for 15 min at 4 °C to obtain liquid mucilage. This mucilage was then freeze-dried for 72 h to form a powder, which was stored in a sealed container in a cool, dry, and dark environment until further use.

Preparation of Pomegranate Peel Extracts (PPE)

The pomegranate peel extract was prepared according to the procedure described by Makhathini, Kaseke and Fawole [36]. Briefly, pomegranate peel powder (300 g) was mixed with 70% ethanol in a 1:3 w/v ratio. The mixture was sonicated in a Labotec ultrasound bath (Cape Town, Western Cape Province, South Africa) at 0 °C for 2 h at maximum power (700) and frequency (40 kHz). After sonication, the samples were filtered using Whatman filter paper No.1 and then centrifuged in a Thermo Fisher Scientific Biofuge centrifuge at 5590 x g and 4 °C for 5 min. This extraction process was repeated twice. The extracts were then concentrated to 9 ° Brix using a BUCHI Rotavapor R-300 (Flawil, Switzerland) at 50 °C under reduced pressure (150 mbar) and stored at 4 °C with 95% RH until further analysis.

Edible Coating Formulation, Optimisation and Application

To prepare the edible coating, mucilage powder extracted from *Opuntia ficus-indica* cladodes was mixed with distilled water to formulate 3% (w/v) mucilage. Pomegranate peel extract (PPE; 1% (w/w)) and cellulose nanofiber (CNF 0.02% (w/v)) were further added. The mixture was continuously stirred on a magnetic plate set at 30 °C for 4 h and homogenized at 5000 rpm to ensure complete hydration and solubility. Glycerol (0.252% (v/v)) was then added as a plasticizer to the coating solutions. This coating formulation was carried out from our previous study based on the optimisation of CNF and glycerol on OFIM-infused PPE [25]. Furthermore, bananas were coated by dipping them in the optimised formulation (OFFIM) for 3 min, while the control group was dipped in distilled water and then dried at room temperature for approximately 1 h. Both coated and control groups were packed in ventilated plastic bags to simulate retail conditions, with each treatment having 3 replicates and each replicate consisting of 7 bananas per storage interval (21 bananas per treatment per storage interval). Bananas were then subjected to retail conditions (23 ± 2 °C with $85 \pm 2\%$ relative humidity).

Fruit Weight Loss

Seven randomly selected bananas per replicate per treatment were weighed individually at the start of the storage experiment using an electronic scale, with initial weights recorded as W_i . The same bananas were then weighed every two days over a ten-day period, with each subsequent weight recorded as W_f . The percentage of weight loss (WL%) was calculated using the Eq. (1).

$$WL (\%) = \frac{W_i - W_f}{W_i} \times 100 \quad (1)$$

Peel Colour Attributes

Banana colour attributes were measured using a colourimeter in CIELAB coordinates (a^* and b^*). Before each reading, the colourimeter was calibrated with a white calibration tile. At each sampling interval, colour measurements were taken from seven bananas per replication (two readings per fruit, once on each opposite side), amounting to 21 bananas per treatment. For each fruit, readings were taken on two sides, and the results were averaged per fruit.

Ethylene Production and Respiration Rate

The respiration rate and ethylene production were measured using the closed system method as Fawole and Opara [37] outlined. In brief, seven bananas were placed in a 3 L glass jar, sealed with a rubber septum lid, and kept at room temperature for 1 h. The carbon dioxide (CO₂) production inside the jar was measured afterward using an infrared gas analyser. The respiration rate was expressed as mL CO₂/(kg.h). Ethylene production was measured with an SCS56 ethylene analyser, with results reported as $\mu\text{L C}_2\text{H}_4/(\text{kg.h})$.

Firmness

Banana firmness was evaluated using a texture analyser fitted with a 5 mm diameter cylindrical probe. The probe penetrated the banana pulp to a depth of 5 mm at a speed of 1.0 mm/s. Measurements were taken around the central zone of banana pulp, and the maximum was recorded in Newtons (N).

Data Analysis

The data were subjected to statistical analysis using R with associated libraries to highlight the effectiveness of the optimised edible coating on fruit quality attributes. Statistical mean differences were shown using Student's t-test at the 5% significance level. All experiments were performed in triplicate, and all physicochemical analysis data are expressed as the mean $\pm$ standard error.

Machine Learning Applications

Workflow Diagram

The workflow for predictive modeling of banana quality using machine learning algorithms is shown in Fig. 1. The process begins with the original dataset of banana quality metrics, followed by data pre-processing to prepare the data for analysis. An exploratory data analysis was then conducted to understand the patterns and relationships within the data. The dataset is then split into training and testing sets (70/30). The training dataset is used to train the model, while the test dataset is used to evaluate the performance of the model. In this study, three machine learning algorithms (Partial Least Squares, Elastic Net regression, and Ridge regression) were applied to the training data.

Data Preprocessing

This operation included data cleaning and transformation. Hence, data were handled to identify and treat the missing

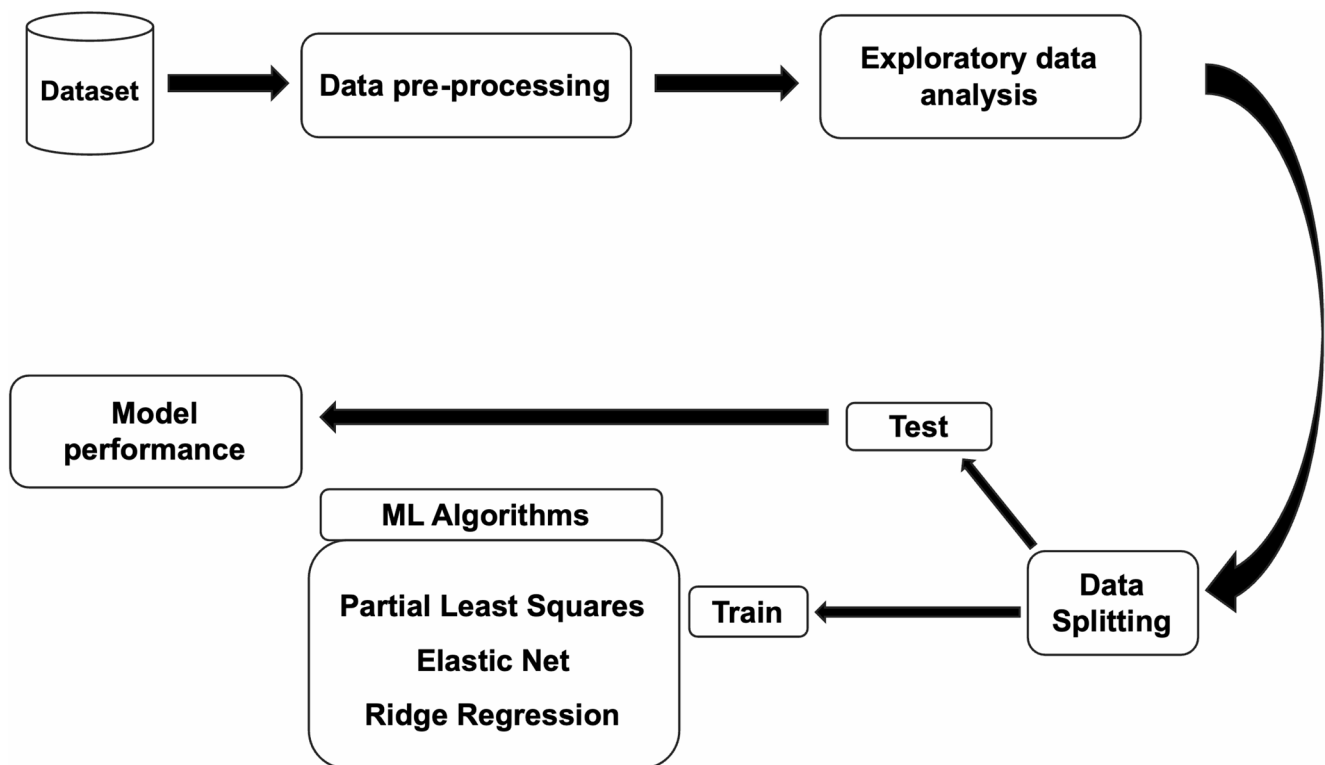


Fig. 1 The workflow for predictive modeling of banana quality using machine learning algorithms

data and outliers. The predictors used in this study were storage time (X_1), weight loss (X_2), respiration rate (X_3), ethylene production (X_4), greenness (X_5), and yellowness (X_6). Given that these features have different units, data were normalised for models considering multiple features to affect them the same importance, using the Minmax normalisation method according to the following Eq. (2) [38].

$$x_N = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (2)$$

Models' Implementation for Predicting Firmness

Simple Linear Regression

Simple linear regression uses the equations for a line to model the relationship between variables. If y is the outcome variable and x is the predictor variable, then

$$Y = kx + c \quad (3)$$

where k is a coefficient that represents the slope of the linear relationship between the variables x and y , and c is a constant.

Partial Least Squares (PLS)

PLS regression was employed to model the relationship between non-invasive parameters and firmness, particularly due to multicollinearity among predictors. PLS regression offers a solution by reducing the predictor variables into a smaller set of uncorrelated latent variables, while preserving predictive accuracy. The model was optimised by tuning the number of latent components ($n_components$) within the range from 1 to $n_features$ (number of features) using GridSearchCV of Scikit learn in Python. A 5-fold cross validation is applied to minimise the risk of overfitting, and the root mean square error (RMSE) was used to identify the optimal number of components across the validation fold.

Elastic Net Regression

Elastic Net regression was used to model the relationship between non-invasive attributes and firmness due to its ability to handle multicollinearity and perform variable selection. By applying Lasso and Ridge penalties Elastic net offers a balance between feature selection and coefficient shrinkage, making it well-suited for datasets with highly correlated predictors. GridSearchCV from the scikit-learn library in Python was employed to optimise hyperparameters of the model. The grid search explores an array of values for regularisation strength α (0.1, 1, 10) and the

L1 ratio (0.1, 0.5, 0.7, 0.9, 0.95, 0.99, 1.0). A 5-fold cross validation is used to prevent overfitting and ensure the model's robustness. The root mean square error (RMSE) was used on validation to identify the optimal combination of parameters.

Ridge Regression

Ridge regression was applied to model the relationship between non-destructive attributes and firmness, primarily to address multicollinearity among predictors. By introducing an L2 regularization term, Ridge regression shrinks correlated coefficients, stabilising the model. The Ridge regression model was optimised by tuning the regularisation parameter (alpha) using GridSearchCV. An array of values was tested (0.05, 1, 10, 100, 100), and the model performance was evaluated through cross validation technique (5 sets). The root mean square error (RMSE) was used to select the optimal alpha.

Evaluation of Models' Performances

The performance metrics used to assess the models' performances in this study included R-squared (R^2), Root Mean Square Error (RMSE), Mean Squared Error (MSE), and R^2 -adjusted.

R-Squared (R^2)

The R-squared was calculated using Eq. 4 to express the proportion of the total variance in the experimental data that can be explained by the model. It varies from 0 to 1, with high values indicating a better fit [39].

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

y_i = Actual value of the dependent variable
 $\hat{y}_i$ = Predicted value of the dependent variable
 $\bar{y}$ = Mean of the actual dependent variable values
 n = Number of observations.

Root Mean Square Error (RMSE)

The RMSE measures the average magnitude of the errors between predicted and actual values. It measures the accuracy of the model by taking the square root of the average of squared differences between predictions and actual observations. It is given by the Eq. (5), according to Wengang et al. [40].

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

y_i = Actual value of the dependent variable
 $\hat{y}_i$ = Predicted value of the dependent variable
 n = Number of observations.

Mean Squared Error (MSE)

MSE represents the average of squared errors between the predicted and actual values. It is also a reliable metric since, like RMSE, it penalizes larger errors than smaller ones. However, it is difficult to read in the original data units without the square root. MSE is calculated using the Eq. (6), described by Chicco et al. [41]

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (6)$$

y_i = Actual value of the dependent variable
 $\hat{y}_i$ = Predicted value of the dependent variable
 n = Number of observations.

R^2 -Adjusted

Adjusted R-squared adjusts the R^2 value according to the number of predictors in the model. It is particularly useful for comparing models with different numbers of features. It is obtained using the Eq. (7).

$$R_{adj}^2 = 1 - \left(\frac{(1 - R^2)(n - 1)}{n - p - 1} \right) \quad (10)$$

R^2 = R-squared of the model.
 n = Number of observations.
 p = number of independent variables (predictors).

Results and Discussion

Weight Loss and Firmness

Weight loss is a crucial indicator of fruit quality and shelf life, as it reflects moisture loss due to respiration and transpiration [42, 43]. Excessive weight loss leads to shrivelling, reduced consumer appeal, and faster senescence, making it a key post-harvest parameter to monitor. Figure 2A illustrates the weight loss trends for control (uncoated) and OF-coated bananas throughout the storage period. The results showed that the weight loss increased progressively in both treatments as storage time advanced. However, a significantly higher weight loss ($p < 0.001$) was observed in the control

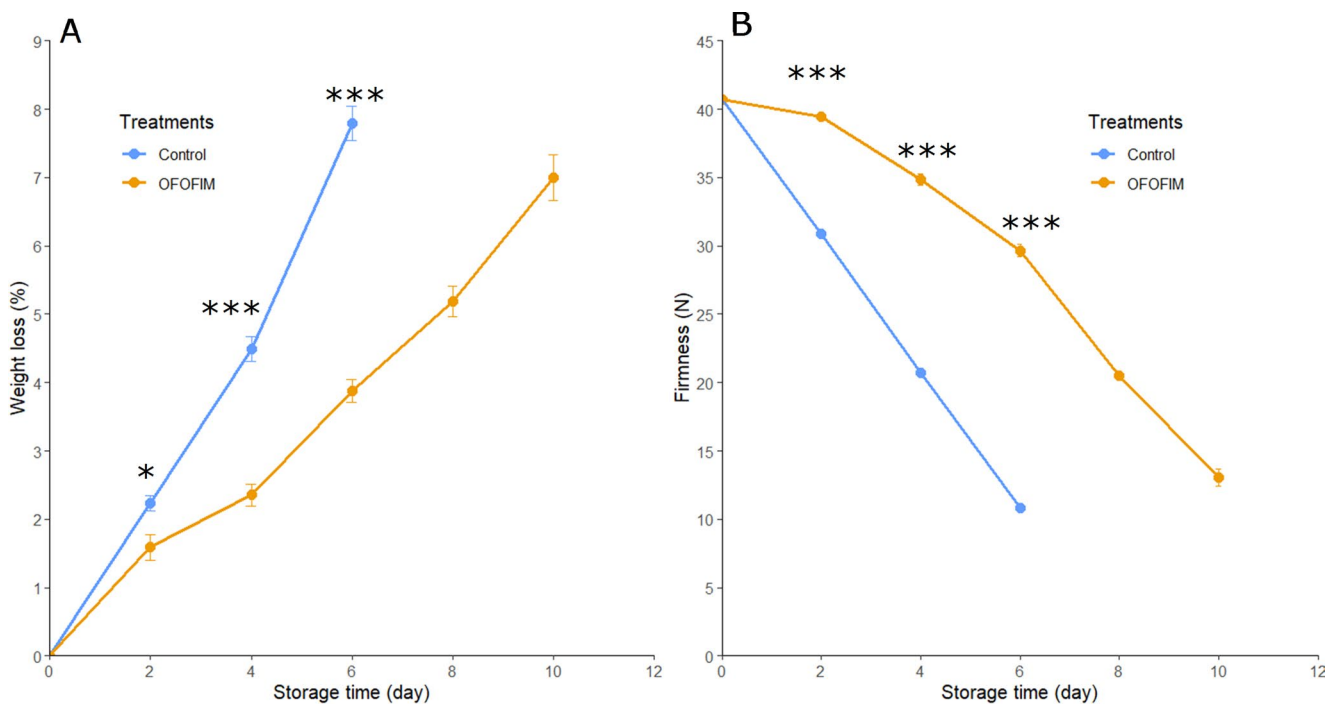


Fig. 2 Changes in weight loss (A) and firmness (B) of bananas over a 10-day storage period at 23 ± 2 °C and $85 \pm 5\%$ RH. OF: optimal formulation of *Opuntia ficus-indica* mucilage. Asterisks indicate statisti-

cally significant differences between control and OF-treated bananas at each time point: *** denotes $p < 0.001$ (highly significant), while * denotes $0.01 \leq p < 0.05$ (significant)

bananas than in the coated fruits, particularly after four days of storage. The superior performance of the OF coating in minimizing weight loss can be attributed to its barrier-forming properties, which effectively reduce moisture evaporation by slowing transpiration and respiration rates [25]. This delayed dehydration and maintained the integrity of the fruit's peel, contributing to an extended shelf life. Additionally, the reduction in weight loss suggests that the coating formulation created a semi-permeable layer that regulated gas exchange, thereby preserving the freshness and marketability of bananas for a longer duration.

Firmness is another critical postharvest attribute influencing consumer perception, transportation stability, and overall fruit quality [29, 44, 45]. Figure 2B shows the changes in firmness over the 10-day storage period. Firmness loss is closely linked to ripening, during which cell wall degradation, enzymatic activity, and water loss contribute to a softer texture [46, 47]. The results indicate that control fruits experienced a rapid decline in firmness, dropping from 40 N to approximately 10 N within six days, whereas OF-coated bananas retained significantly higher firmness throughout the storage period. Statistical differences were observed between the treatments from 2 d to 6 d ($p < 0.001$), highlighting the strong protective effects of the edible coating in delaying firmness loss and preserving textural integrity.

The coating likely preserved firmness by modulating respiration rate, ethylene production, and moisture retention,

all of which contribute to the softening of fruits during ripening [48]. By mitigating these physiological changes, the coating effectively maintained the fruit's structural integrity and prevented premature softening. This finding aligns with previous studies that demonstrated the efficacy of edible coatings in delaying ripening-related changes in tropical fruits [14, 24].

Banana Peel Colour

The greenness parameter (a^*) is an important indicator of banana ripening, as the fruit transitions from green to yellow due to chlorophyll degradation and carotenoid accumulation [49]. As ripening progresses, the a^* value increases, reflecting the loss of green pigments [50, 51]. Figure 3A illustrates that control bananas exhibited a rapid decline in greenness, with a^* values becoming positive after six days of storage, indicating full ripening and complete chlorophyll breakdown.

In contrast, the treated bananas-maintained fruit greenness for a significantly longer period ($p < 0.001$), suggesting a delayed transition in a^* values from day 2 to day 6. This suggests that the coating markedly slowed down chlorophyll degradation, preserving the green colour of bananas and delaying visual ripening symptoms. The preservation of greenness indicates that the coating's barrier properties reduced gaseous exchange, particularly oxygen exposure,

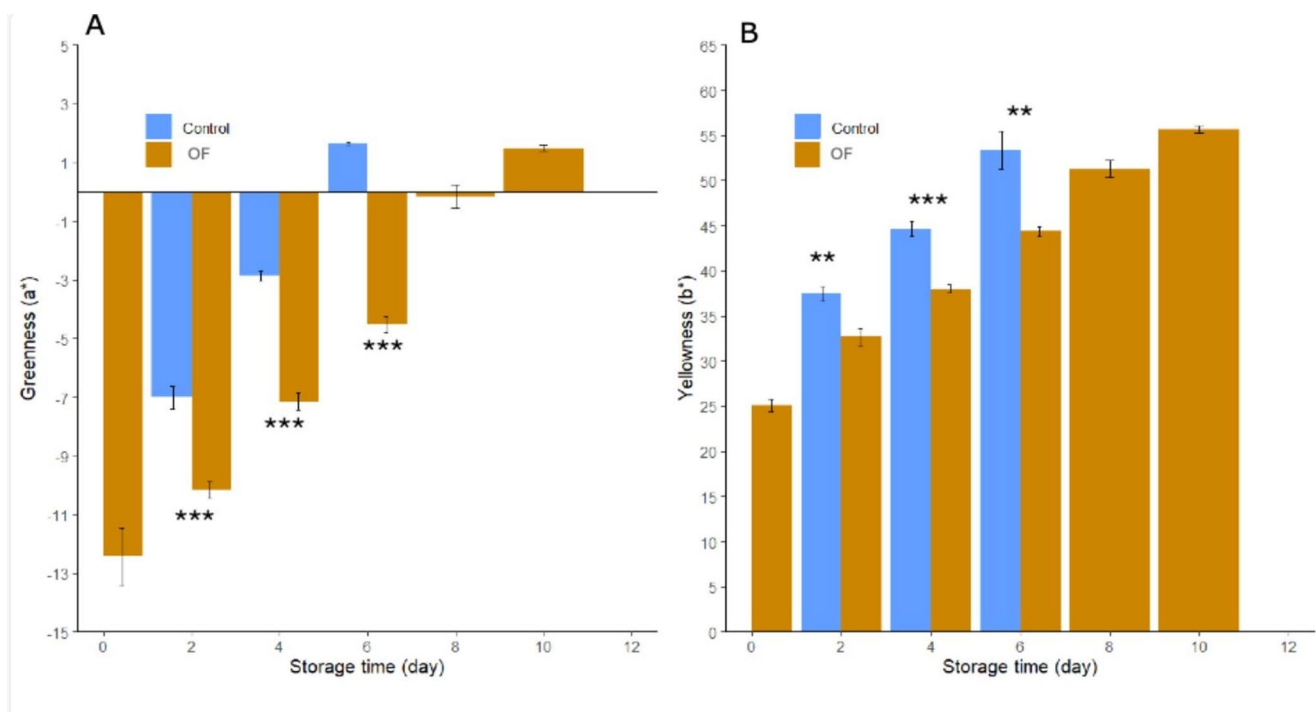


Fig. 3 Changes in greenness (a*) (**A**) and yellowness (b*) (**B**) of bananas over a 10-day storage period at 23 ± 2 °C and $85 \pm 5\%$ RH. OF: optimal formulation of *Opuntia ficus-indica* mucilage. Asterisks

indicate statistically significant differences between control and OF-treated bananas at each time point: *** denotes $p < 0.001$ (highly significant), while * denotes $0.01 \leq p < 0.05$ (significant)

which is known to accelerate oxidative degradation of chlorophyll. These findings align with previous studies reporting that edible coatings can delay chlorophyll breakdown by modulating oxygen and ethylene exchange [13, 48, 52]. Alongside chlorophyll degradation, the development of yellow pigmentation (b*) is another key ripening indicator, reflecting the synthesis of carotenoids [50]. As shown in Fig. 3B, yellowness (b*) progressively increased in control and coated bananas, but the coated bananas exhibited a significantly slower increase in b* values compared to the control at all storage intervals ($p < 0.001$). This suggests that the mucilage-based coating effectively slowed the unmasking of yellow carotenoids, thereby delaying the ripening process.

The delayed colour changes observed in coated bananas can be attributed to the coating's ability to regulate gaseous exchange, particularly ethylene production, which is a key driver of colour transformation during banana ripening [13, 52]. Ethylene is known to activate chlorophyll degradation through chlorophyllase enzyme activity, accelerating the yellowing process [13]. Moreover, previous studies suggest that chlorophyll degradation and pectin solubilization occur concurrently, as both processes are ethylene-dependent [51, 53]. This is consistent with the present study, where bananas that exhibited lower greenness loss (delayed increase in a* values) also maintained higher firmness over time.

Respiration Rate and Ethylene Production

The respiration rate is a fundamental metabolic process of fruit, which generally increases with ripening, leading to faster senescence and spoilage [54]. Therefore, controlling the respiration rate is critical for extending fruit shelf life. In this study, the respiration rate increased progressively in both treatments over the 10-day storage period (Fig. 4A). However, control fruits exhibited a significantly steeper increase in respiration rate compared to coated fruits ($p < 0.001$), indicating accelerated metabolic activity and ripening in the uncoated samples. This suggests that the OF coating effectively delayed ripening by modifying the fruit's immediate atmospheric conditions. Edible coatings are known to create a semi-permeable barrier that regulates gas exchange, reducing oxygen availability while increasing carbon dioxide levels. This modified microenvironment slows down enzymatic activities associated with respiration and ripening [48]. The slower respiration rate observed in coated fruits is, therefore, likely due to reduced oxygen diffusion and a suppressed metabolic response, effectively prolonging freshness and delaying senescence.

Ethylene production is another key factor influencing fruit ripening, as this plant hormone serves as a ripening trigger in climacteric fruits such as bananas [55]. Elevated ethylene levels accelerate physiological changes, including increased respiration, firmness loss, starch-to-sugar conversion, and

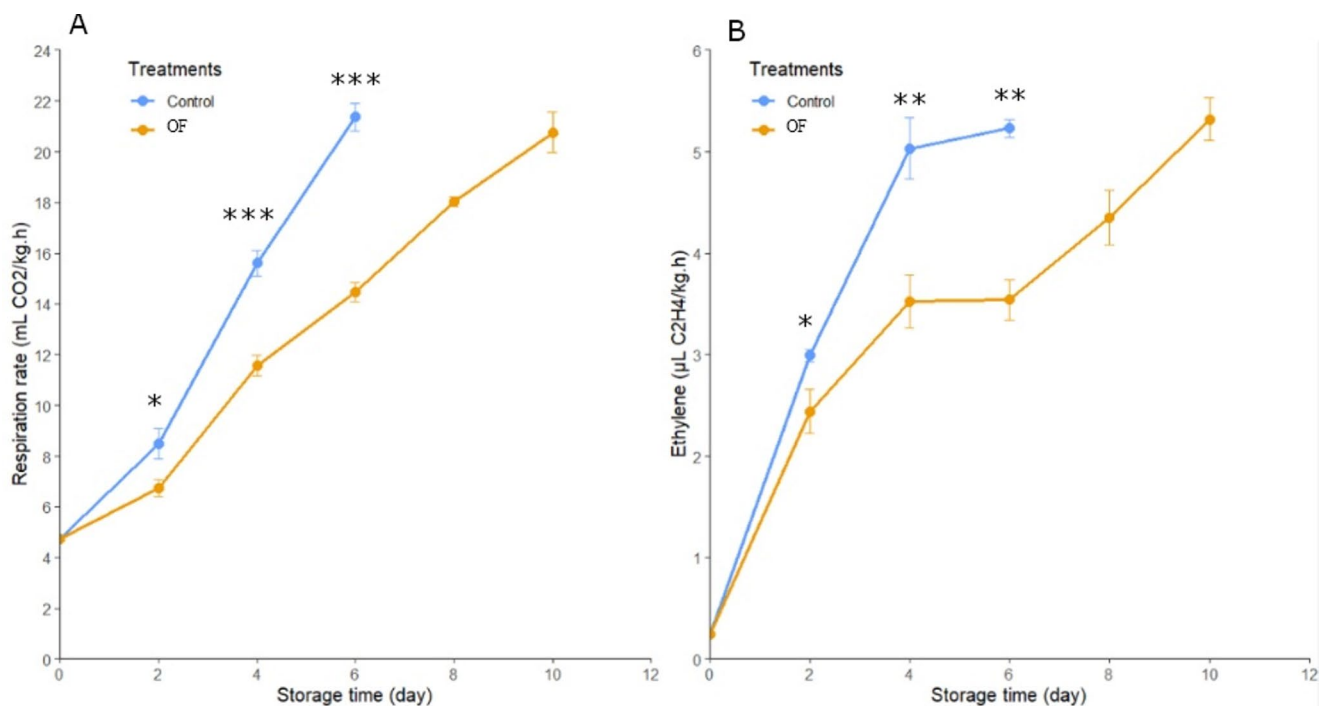


Fig. 4 Changes in respiration rate (**A**) and ethylene production (**B**) of bananas over a 10-day storage period at 23 ± 2 °C and $85 \pm 5\%$ RH. OF: optimal formulation of *Opuntia ficus-indica* mucilage. Asterisks

indicate statistically significant differences between control and OF-treated bananas at each time point: *** denotes $p < 0.001$ (highly significant), while * denotes $0.01 \leq p < 0.05$ (significant)

colour transitions [56]. While ethylene-induced changes can enhance aroma and sugar accumulation, they also promote rapid softening and increased susceptibility to microbial decay. Thus, controlling ethylene production is essential for extending storage life and maintaining fruit quality during postharvest handling. Figure 4B illustrates the ethylene production trends over 10 days of storage. Both control and coated bananas exhibited increased ethylene levels, however, the control fruits showed a significantly higher ethylene peak from day 2 to day 6 ($p < 0.001$), correlating with accelerated ripening. In contrast, coated bananas maintained lower ethylene production throughout storage, indicating that the OF coating effectively suppressed ethylene biosynthesis or limited its diffusion. These findings align with previous studies, which report that edible coatings reduce ethylene biosynthesis by limiting oxygen availability and decreasing internal ethylene concentrations [13, 24]. By lowering both respiration and ethylene production rates, the mucilage-based coating delayed the metabolic processes responsible for fruit softening, weight loss, and colour changes.

Features Correlation

Since the quality attributes investigated in this study are closely associated with the ripening process, understanding their interrelationships provides critical insights into the mechanisms governing banana firmness loss. To explore

these underlying patterns, we conducted a correlation analysis between firmness and key non-destructive predictors. The results, presented in Fig. 5, demonstrated the strong linear relationships between all measured parameters. All predictors, including weight loss, respiration rate, ethylene production, and colour attributes (greenness and yellowness), exhibited strong positive correlations with each other while showing negative correlations with firmness, the target variable. This trend confirms that as bananas ripen, respiration rate, ethylene production, weight loss, and yellowness (b) increase, whereas firmness and greenness (a) progressively decline [13, 24, 48]. The strongest correlations were observed between firmness and weight loss (-0.97), respiration (-0.95), greenness (-0.96), and yellowness (-0.94), indicating that these non-destructive indicators are highly predictive of textural changes during ripening.

These findings support the feasibility of using non-destructive quality attributes for predicting banana firmness with high accuracy. The observed correlations align with previous studies, which highlight the role of respiration rate and ethylene production in accelerating ripening-related changes in firmness and colour [13, 24]. Additionally, the presence of strong correlations suggests a degree of multicollinearity among the predictors, which influences the selection of suitable machine-learning models for predictive analysis. The scatterplot matrix in Fig. 5 highlights the interdependence of these parameters and their significant

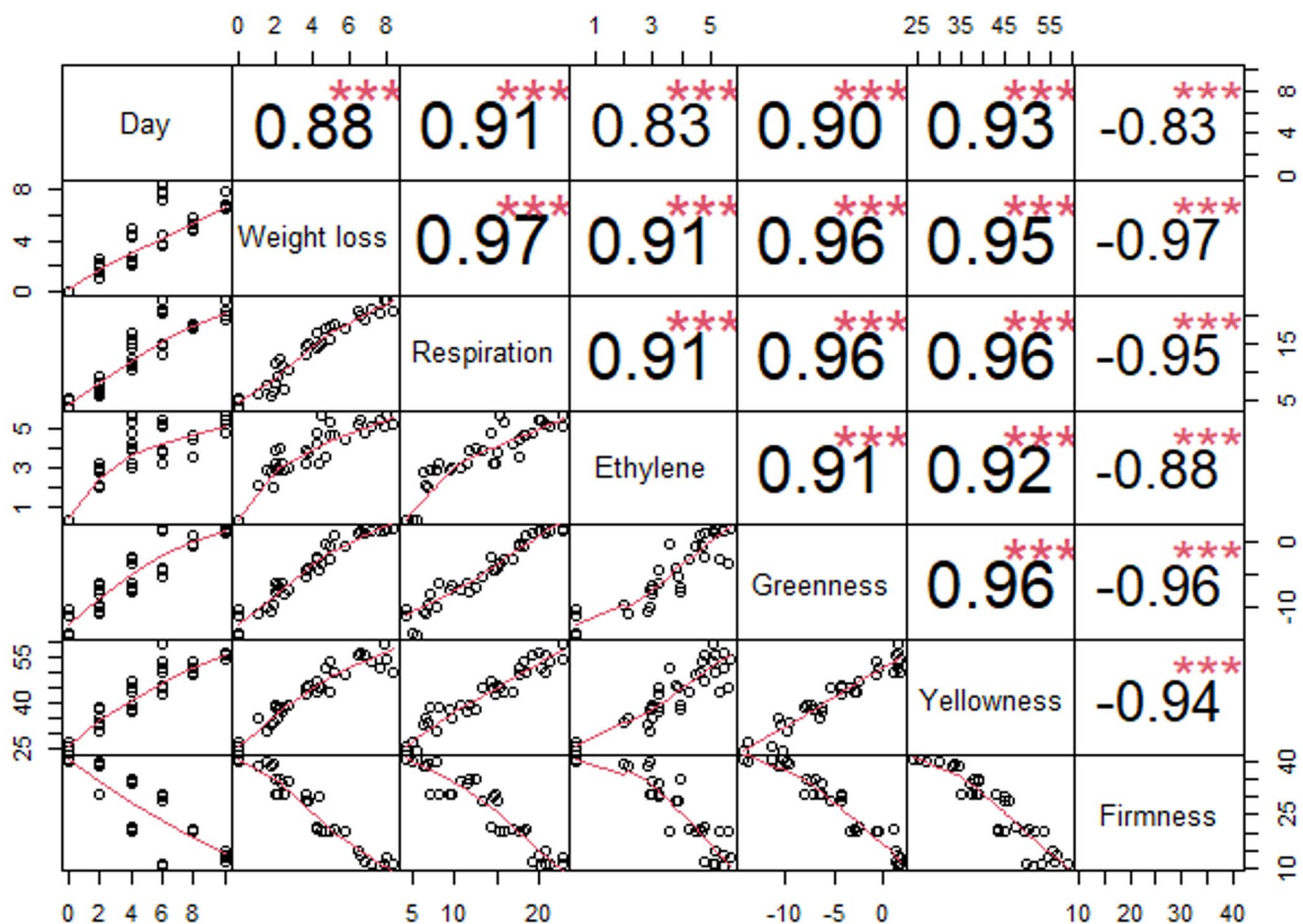


Fig. 5 Correlation matrix showing the relationships between firmness and non-destructive parameters (days, weight loss, respiration rate, ethylene production, greenness, and yellowness) over a 10-day storage

period at 23 ± 2 °C and $85 \pm 5\%$ RH for control and coated bananas. The asterisks indicate significance levels: *** ($p < 0.001$; highly significant)

influence on banana firmness over time. This further justifies the use of multivariate regression-based models, such as Partial Least Squares (PLS) regression, Ridge regression, and Elastic Net regression, which appropriately handle collinear features in predictive modelling.

Regression Plot

Figure 6 presents the regression plots depicting the relationship between firmness and weight loss, respiration rate, ethylene production, greenness, and yellowness, across the 10-day storage period for both control and OF-coated bananas. Firmness declined progressively in both treatments, with a steeper reduction in the control group (slope = -4.98) compared to coated bananas (slope = -2.86), indicating the coating's effectiveness in delaying softening by minimising moisture loss and metabolomic activity. Previous studies have suggested that edible coatings act as barriers to water vapour, reducing transpiration, subsequent water loss

from the fruit. This directly helps in retaining fruit firmness, as dehydration is a major cause of texture decline [57, 58].

Respiration rate also correlated with firmness loss, with slightly steeper slopes observed in the control (-1.70) versus coated (-1.66) samples, suggesting that the coating moderated respiration-induced softening. In contrast, ethylene production showed nearly identical slopes for both treatments (control = -5.15; coated = -5.20), underscoring its central role in ripening and cell wall breakdown. This is consistent with a high negative correlation ($r = -0.88$, Fig. 5). Prior studies have suggested that edible coatings slow respiration and inhibit ethylene synthesis, both of which are key drivers of ripening and softening in climacteric fruit [59, 60].

Colour changes were similarly associated with firmness decline. The loss of greenness (a^*) showed a stronger impact in the control group (slope = -2.08) than in coated fruit (-1.90), suggesting delayed chlorophyll degradation. Increasing yellowness (b^*) was also linked to firmness reduction, reinforcing the connection between visual ripening and texture softening. Predictive equations and

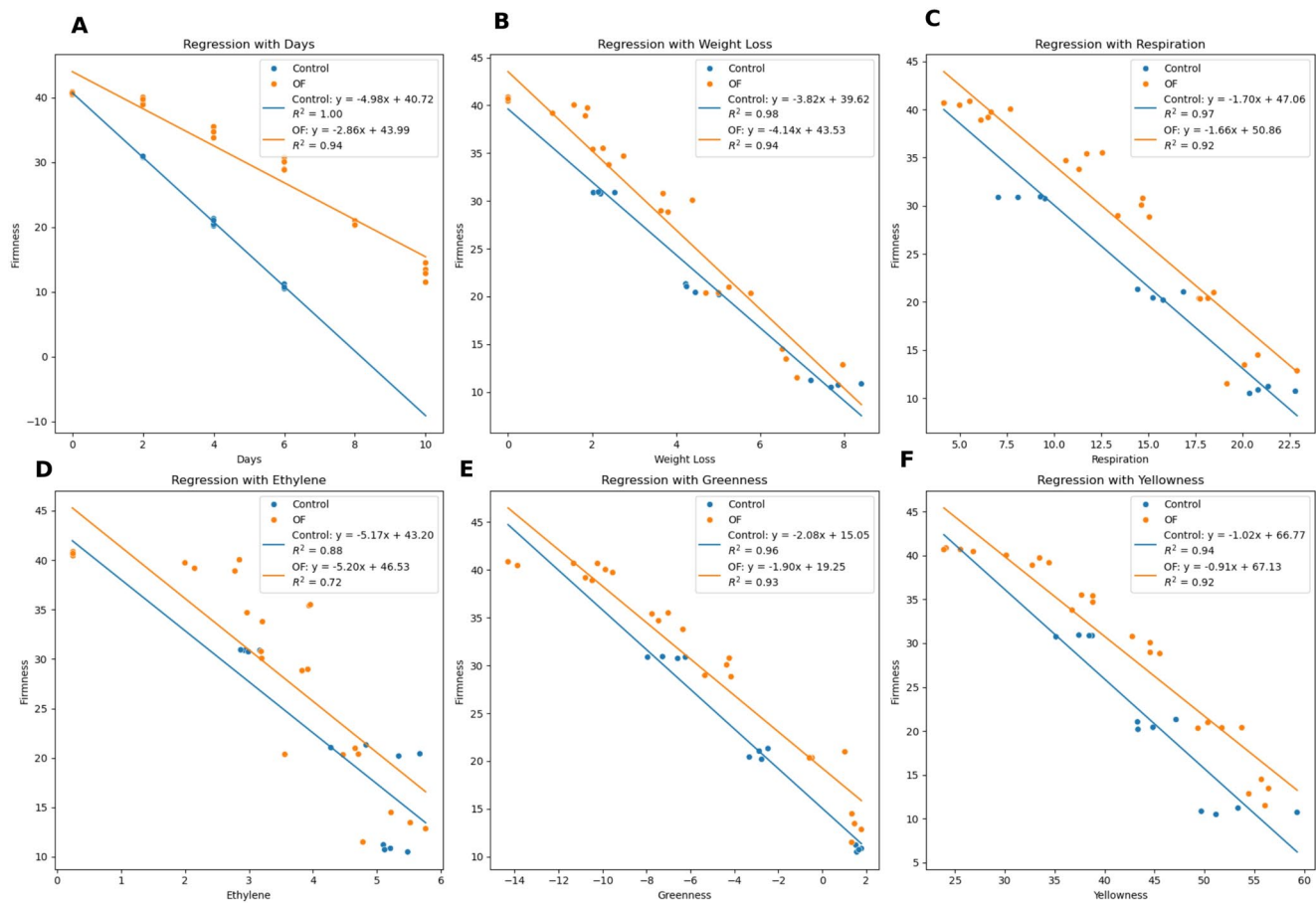


Fig. 6 The relationship between firmness and non-destructive parameters days (A), weight loss (B), respiration rate (C), ethylene production (D), greenness (E), and yellowness (F) of bananas over a 10-day

Table 1 Performance comparison of machine learning algorithms for predicting firmness using regression metrics

Machine Learning Algorithms	Performance metrics			
	R^2	RMSE	MSE	R^2 -adjusted
Ridge Regression	0.972	0.110	0.012	0.922
PLS Regression	0.978	0.097	0.009	0.940
Elastic Net	0.956	0.142	0.020	0.801

R^2 : Coefficient of determination; RMSE: Root Mean Squared Error; MSE: Mean Squared Error; MAE: Mean Absolute Error

performance metrics for firmness estimation based on these attributes are provided in Table S1 (Supplementary Data).

Multiple Linear Regression

Table 1 presents the performance metrics of three machine learning algorithms, including Ridge regression, Partial Least Square (PLS) Regression, and Elastic Net Regression. The best-performing model was selected based on the coefficient of determination (R^2), Root Mean Square Error (RMSE), Mean Squared Error (MSE), and R^2 -adjusted. Among the tested models, PLS Regression demonstrated the

storage period at 23 ± 2 °C and $85 \pm 5\%$ RH for control and optimal formulation of *Opuntia ficus-indica* mucilage (OF)

highest predictive accuracy, achieving an R^2 value of 0.978. This indicates that 97.8% of the variance in banana firmness was explained by the model, making it slightly superior to Ridge Regression ($R^2 = 0.972$). Elastic Net Regression exhibited the lowest R^2 value of 0.956, accounting for 95.6% of the variance in the dependent variable. These findings confirm that PLS Regression outperformed the other models across all performance metrics, highlighting its suitability for firmness prediction. The robustness of PLS regression is likely attributed to its ability to handle multicollinearity among predictors while capturing latent structures that explain variance in both independent and dependent variables. This feature makes it particularly effective when dealing with datasets containing highly correlated predictors, as observed in this study. Moreover, Given the complex and potentially noisy nature of physicochemical data collected, PLS's dimensionality reduction and emphasis on covariance likely contributed to its effectiveness.

The superior predictive performance observed in this study aligns with previous research utilizing PLS Regression for fruit firmness prediction. For instance, Dias et al.,

[28] applied PLS Regression to develop a mathematical model for predicting pear firmness, which accounted for 89% of the variance in firmness. This prediction accuracy was lower than that achieved in the current study ($R^2 = 0.978$), suggesting that the selected non-destructive predictors provided a more comprehensive assessment of firmness in bananas. Similarly, a study by Ivanovski et al., [61] evaluated different machine learning algorithms for predicting peach firmness, reporting an MSE value of 0.67 for simple linear regression, 0.63 for multiple linear regression, and 0.65 for artificial neural networks. In contrast, our study achieved a significantly lower MSE value, with the highest observed at 0.142, further demonstrating the effectiveness of the selected machine learning approaches. Additionally, Ljubobratović et al., [62] used a black-box machine learning approach to predict apple firmness, achieving an accuracy of 97.5% on test data, which is closely comparable to the performance observed in this study. Chen et al., [63] achieved high accuracy in predicting the firmness of yellow flesh peach firmness by integrating acoustic and Vis/NIR spectroscopy techniques, reporting R^2 of 0.951 and RMSE of 0.443. Moreover, a study by Wang, Feng, Ji and Cui [64] employed an excitation source combined with a modified one-dimensional convolutional neural network (1D-CNNm) algorithms to predict peach firmness, yielding only moderate predictive performance, with R^2 of 0.794.

The high predictive accuracy achieved in this study highlights the potential of machine learning-based models for non-invasive firmness prediction in bananas. This study demonstrated the practical application of machine learning models to accurately predict banana firmness under retail conditions using an optimised edible coating. These findings not only advance the scientific understanding of postharvest preservation but also pave the way for industrial-scale integration of intelligent quality monitoring systems. The food industry could benefit from predictive modelling to enhance shelf-life management and reduce postharvest losses.

Conclusion

This study demonstrated that the OF effectively delayed banana fruit ripening, extending the shelf life by four days. All non-destructive attributes exhibited strong positive correlations with each other while showing an inverse relationship with firmness. These findings confirm that these parameters are reliable indicators of firmness and can be effectively used for the non-destructive prediction of banana ripeness. Among the machine learning models tested, PLS Regression emerged as the best-performing model. Ridge Regression and Elastic Net Regression also showed good performance, however, PLS outperformed

them overall. These findings highlight the potential of combining optimised edible coatings with machine learning models to enhance fruit shelf life, improve quality control, and reduce postharvest losses in the banana industry. The ability to predict firmness using non-destructive attributes offers a practical, real-time solution for postharvest management, supporting data-driven decisions to maintain fruit freshness, reduce waste, and optimise supply chain logistics. Moreover, the natural biobased coating materials used in this study hold potential for adaptation in cosmetic and pharmaceutical applications, such as biodegradable packaging, controlled release systems. Although these models predict accurately the banana firmness using non-destructive attributes, some limitations must be acknowledged. The relatively small data might reduce the generalisation of our findings. In addition, the controlled laboratory conditions may not completely reflect the complexity of real-world (packhouse) and environmental conditions. Therefore, to provide a practical and real-time solution for postharvest management of bananas, further research could consider a larger dataset across multiple cultivars and seasons collected from real-world supply chains.

Acknowledgements This work is based on research supported by the National Research Foundation of South Africa (SPAR231013155231) and the University Research Committee at the University of Johannesburg, South Africa.

Author Contributions Mawande Shinga: Writing—original draft, investigation, Visualization, Methodology, Formal analysis, Data curation, Software. Yadjouma Silue: Review & editing, Data curation, Software, Conceptualization. Olaniyi Fawole: Review & editing, Conceptualization, Supervising, Project administration, Funding acquisition.

Funding Open access funding provided by University of Johannesburg.

Data Availability No datasets were generated or analysed during the current study.

Declarations

Competing Interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Z. Zhai, C. Feng, Y. Wang, Y. Sun, X. Peng, Y. Xiao, X. Zhang, X. Zhou, J. Jiao, W. Wang, *Int. J. Mol. Sci.* **22**(22), 12331 (2021). <https://doi.org/10.3390/ijms222212331>
- Y. Zheng, L. Duan, Y. Ran, P. Zhang, Y. Jiang, Z. Zhao, Z. Li, L. Chen, Y. Tang, X. Li, *Postharvest Biol. Technol.* **196**, 112190 (2023). <https://doi.org/10.1016/j.postharvbio.2022.112190>
- W. Yan, M. Cao, L. Shi, W. Wu, F. Xu, W. Chen, Z. Yang, *Postharvest Biol. Technol.* **213**, 112916 (2024). <https://doi.org/10.1016/j.postharvbio.2024.112916>
- S. Forlani, S. Masiero, C. Mizzotti, *J. Exp. Bot.* **70**(11), 2993–3006 (2019). <https://doi.org/10.1093/jxb/erz112>
- S. Maduwanthi, R. Marapana, *International journal of food science* 2019 (1), 2520179 (2019). <https://doi.org/10.1155/2019/2520179Citations>
- R.B. Watharkar, S. Chakraborty, P.P. Srivastav, B. Srivastava, *J. Food Meas. Charact.* **15**, 3336–3349 (2021). <https://doi.org/10.1007/s11694-021-00907-9>
- M. Hailu, T. Seyoum Workneh, D. Belew, *J. Food Sci. Technol.* **51**, 2947–2963 (2014). <https://doi.org/10.1007/s13197-012-0826-5>
- S.B. Murmu, H.N. Mishra, *Innovative Food Sci. Emerg. Technol.* **49**, 20–30 (2018). <https://doi.org/10.1016/j.ifset.2018.07.011>
- J. Wei, F. Ma, S. Shi, X. Qi, X. Zhu, J. Yuan, *Postharvest Biol. Technol.* **56**(2), 147–154 (2010). <https://doi.org/10.1016/j.postharvbio.2009.12.003>
- S. Musacchi, S. Serra, *Sci. Hort.* **234**, 409–430 (2018). <https://doi.org/10.1016/j.scienta.2017.12.057>
- V. Paul, R. Pandey, *J. Food Sci. Technol.* **51**, 1223–1250 (2014). <https://doi.org/10.1007/s13197-011-0583-x>
- H. Chen, S. Cao, X. Fang, H. Mu, H. Yang, X. Wang, Q. Xu, H. Gao, *Sci. Hort.* **188**, 44–48 (2015). <https://doi.org/10.1016/j.scienta.2015.03.018>
- S. Ali, M.A. Ullah, A. Nawaz, S. Naz, A.A. Shah, G. Gohari, F. Razavi, G. Khaliq, K. Razzaq, *Postharvest Biol. Technol.* **191** (2022). <https://doi.org/10.1016/j.postharvbio.2022.111978>
- S. Ali, N. Zahid, A. Nawaz, S. Naz, S. Ejaz, S. Ullah, B. Siddiq, *Int. J. Biol. Macromol.* **222** (Pt A), 521–532 (2022). <https://doi.org/10.1016/j.ijbiomac.2022.09.159>
- X. Li, T. Xiong, Q. Zhu, Y. Zhou, Q. Lei, H. Lu, W. Chen, X. Li, X. Zhu, *Postharvest Biol. Technol.* **198** (2023). <https://doi.org/10.1016/j.postharvbio.2023.112265>
- T. Wang, Y. Song, L. Lai, D. Fang, W. Li, F. Cao, E. Su, *Food Packaging Shelf Life.* **46** (2024). <https://doi.org/10.1016/j.fpsl.2024.101386>
- Y. Wang, Z. Luo, Z.U. Khan, L. Mao, T. Ying, *Postharvest Biol. Technol.* **108**, 21–27 (2015). <https://doi.org/10.1016/j.postharvbio.2015.05.007>
- O. Khademi, M. Ashtari, F. Razavi, *Sci. Hort.* **249**, 334–339 (2019). <https://doi.org/10.1016/j.scienta.2019.02.018>
- A.F. Ali, A.A. Hatamnia, P. Malekzadeh, M. Sayyari, M.S. Aghdam, *Postharvest Biol. Technol.* **205**, 112521 (2023). <https://doi.org/10.1016/j.postharvbio.2023.112521>
- N. Niaz, W. Hu, S. Pan, K. Ali, S. Niaz, *Postharvest Biol. Technol.* **216**, 113077 (2024). <https://doi.org/10.1016/j.postharvbio.2024.113077>
- A. Allegra, P. Inglese, G. Sortino, L. Settanni, A. Todaro, G. Liguori, *Postharvest Biol. Technol.* **120**, 45–51 (2016). <https://doi.org/10.1016/j.postharvbio.2016.05.011>
- M.A. Awad, A.D. Al-Qurashi, S.A. Mohamed, R.M. El-Shishtawy, M.A. Ali, *J. Food Sci. Technol.* **54**(2), 447–457 (2017). <https://doi.org/10.1007/s13197-016-2481-8>
- M. Shinga, A. Mditshwa, L. Magwaza, S. Tesfay, presented at the II International Symposium on Moringa 1306, 2019. <https://doi.org/10.17660/ActaHortic.2021.1306.38>
- M.H. Shinga, O.A. Fawole, *Int. J. Biol. Macromol.* **245**, 125550 (2023). <https://doi.org/10.1016/j.ijbiomac.2023.125550>
- M.H. Shinga, T. Kaseke, T.M. Pfuakwa, O.A. Fawole, *Food Packaging Shelf Life.* **47**, 101428 (2025). <https://doi.org/10.1016/j.fpsl.2024.101428>
- Y. Huang, R. Lu, K. Chen, *J. Food Eng.* **222**, 185–198 (2018). <https://doi.org/10.1016/j.jfoodeng.2017.11.030>
- B. Yang, W. Guo, X. Huang, R. Du, Z. Liu, *Comput. Electron. Agric.* **179**, 105831 (2020). <https://doi.org/10.1016/j.compag.2020.105831>
- C. Dias, T.R. Brandão, C. Cassoni, M.W. Vasconcelos, A. Ferrante, M. Pintado, *Postharvest Biol. Technol.* **205**, 112493 (2023). <https://doi.org/10.1016/j.postharvbio.2023.112493>
- L. Chen, U.L. Opara, *Food Res. Int.* **51**(2), 823–835 (2013). <https://doi.org/10.1016/j.foodres.2013.01.046>
- S.S.A. Shah, A. Zeb, W.S. Qureshi, M. Arslan, A.U. Malik, W. Alasmary, E. Alanazi, *Infrared Phys. Technol.* **111**, 103479 (2020). <https://doi.org/10.1016/j.infrared.2020.103479>
- I.K. Opara, U.L. Opara, J.A. Okolie, O.A. Fawole, *Plants.* **13**(9), 1200 (2024). <https://doi.org/10.3390/plants13091200>
- H. Azarmdel, A. Jahanbakhshi, S.S. Mohtasebi, A.R. Muñoz, *Postharvest Biol. Technol.* **166**, 111201 (2020). <https://doi.org/10.1016/j.postharvbio.2020.111201>
- H.S. El-Mesery, O.A. Adelusi, S. Ghashi, P.B. Njobeh, Z. Hu, W. Kun, *LWT.* **116**, 116280 (2024). <https://doi.org/10.1016/j.lwt.2024.116280>
- M.D. Elbi, E.Ö. Çapraz, E. Şahin, M.U. Koyuncuoğlu, C. Tuncer, *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi.* **30**(3), 343–353 (2023). <https://doi.org/10.5505/pajes.2023.71242>
- I. Ivanova, M. Serdyuk, T. Tymoshchuk, V. Malkina, A. Shkinder-Barmina, A. Drobitko, N. Zahorko, Y. Muliienok, A. Shepel, Y. Savchuk, *Future Food: J. Food Agric. Soc.* **12**(1) (2024). <https://doi.org/10.17170/kobra-202407218427>
- N. Makhathini, T. Kaseke, O.A. Fawole, *J. Agric. Food Res.* **14**, 100780 (2023). <https://doi.org/10.1016/j.jafr.2023.100780>
- O.A. Fawole, U.L. Opara, *Ind. Crops Prod.* **47**, 300–309 (2013). <https://doi.org/10.1016/j.indcrop.2013.03.028>
- C. Yang, W. Jiang, Z. Guo, *IEEE Access.* **7**, 155304–155312 (2019). <https://doi.org/10.1109/ACCESS.2019.2949287>
- I. Gupta, D. Devegoda, V. Jayaram, C. Rai, C. Sondergeld, *Interpretation.* **7**(3), SF41–SF55 (2019)
- H. Wengang, W. Xiyu, M. Jiajie, G. Ping, W. Lei, *Innovative Food Sci. Emerg. Technol.* **97**, 103836 (2024)
- D. Chicco, M.J. Warrens, G. Jurman, *Peerj computer science* **7**, e623 (2021). <https://doi.org/10.7717/peerj-cs.623>
- R. Lufu, A. Ambaw, U.L. Opara, *Postharvest Biol. Technol.* **157**, 110982 (2019). <https://doi.org/10.1016/j.postharvbio.2019.110982>
- R.A. Sanad Alsbu, P. Yarlagadda, A. Karim, *Foods.* **12**(11), 2113 (2023). <https://doi.org/10.3390/foods12112113>
- C. Bugaud, I. Maraval, M.-O. Daribo, N. Leclerc, F. Salmon, *Sci. Hort.* **211**, 399–409 (2016). <https://doi.org/10.1016/j.scienta.2016.09.016>
- A.A.N. Gunny, S.C. Gopinath, A. Ali, C. Wongs-Aree, N.H.M. Salleh, *J. Agric. Food Res.* **101**, 249 (2024). <https://doi.org/10.1016/j.jafr.2024.101249>
- Y.-Ren, P.-Sun, X.-Wang, Z.-Zhu, *Postharvest Biol. Technol.* **166**, 111203 (2020). <https://doi.org/10.1016/j.postharvbio.2020.111203>
- K. Jia, W. Wang, Q. Zhang, W. Jia, *Int. J. Mol. Sci.* **24**(4), 4054 (2023). <https://doi.org/10.3390/ijms24044054>

48. R. Thakur, P. Pristijono, M. Bowyer, S.P. Singh, C.J. Scarlett, C.E. Stathopoulos, Q.V. Vuong, *Lwt.* **100**, 341–347 (2019). <https://doi.org/10.1016/j.lwt.2018.10.055>
49. P.B. Pathare, U.L. Opara, F.A.-J. Al-Said, *Food Bioprocess Technol.* **6**(1), 36–60 (2012). <https://doi.org/10.1007/s11947-012-0867-9>
50. V. Srilaong, S. Aiamla-or, A. Soontornwat, M. Shigyo, N. Yamauchi, *Postharvest Biol. Technol.* **59**(1), 110–112 (2011). <https://doi.org/10.1016/j.postharvbio.2010.07.006>
51. N. Charoenchongsuk, K. Ikeda, A. Itai, A. Oikawa, H. Murayama, *Sci. Hort.* **181**, 89–94 (2015). <https://doi.org/10.1016/j.scienta.2014.10.005>
52. M. Al-Dairi, P.B. Pathare, R. Al-Yahyai, H. Jayasuriya, Z. Al-Attabi, *Trends Food Sci. Technol.* **134**, 177–191 (2023). <https://doi.org/10.1016/j.tifs.2023.03.003>
53. M.M. Tipu, S.M. Sherif, *Front. Plant Sci.* **15**, 1475496 (2024). <https://doi.org/10.3389/fpls.2024.1475496>
54. V. Prasanna, T. Prabha, R. Tharanathan, *Crit. Rev. Food Sci. Nutr.* **47**(1), 1–19 (2007). <https://doi.org/10.1080/10408390600976841>
55. M.R. Alina, C.C. Mureşan, A. Pop, G.S. Martiş, A.E. Mureşan, A.N. Postolache, F. Stoica, I.C. Crivei, I.-D. Veleşcu, R.N. Raţu, (2023). <https://doi.org/10.5772/intechopen.114015>
56. G. Tucker, X. Yin, A. Zhang, M. Wang, Q. Zhu, X. Liu, X. Xie, K. Chen, D. Grierson, *Food Qual. Saf.* **1**(4), 253–267 (2017). <https://doi.org/10.1093/fqsafe/fyx024>
57. P. Kumar, S. Sethi, R. Sharma, M. Srivastav, D. Singh, E. Varghese, *J. Food Sci. Technol.* **55**, 2344–2350 (2018)
58. S. Ngubane, S.Z. Tesfay, L.S. Magwaza, A. Mditshwa, *Front. Sustainable Food Syst.* **8**, 1473731 (2024). <https://doi.org/10.3389/fsufs.2024.1473731>
59. H. Ustun, A. Dogan, B. Peker, C. Ural, M. Cetin, Y. Ozyigit, M. Erkan, *J. Sci. Food Agric.* **103**(1), 176–184 (2023). <https://doi.org/10.1002/jsfa.12129>
60. A.S. Formiga, E.M. Pereira, J.S.P. Junior, F.B. Costa, B.-H. Matiuiz, *Food Chem. Adv.* **1**, 100124 (2022). <https://doi.org/10.1590/1980-5373-mr-2023-0060>
61. T. Ivanovski, G. Zhang, T. Jemrić, M. Gulić, M. Matetić, presented at the 2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO), 2020. <https://doi.org/10.23919/MIPRO48935.2020.9245213>
62. D. Ljubobratović, G. Zhang, M. Brkić Bakarić, T. Jemrić, M. Matetić, presented at the Proceedings of the 31st International DAAAM Symposium, 2020. <https://doi.org/10.2507/31st.daaam.proceedings.099>
63. N. Chen, Z. Liu, T. Zhang, Q. Lai, J. Zhang, X. Wei, Y. Liu, *LWT.* **209**, 116772 (2024). <https://doi.org/10.1016/j.lwt.2024.116772>
64. D. Wang, Z. Feng, S. Ji, D. Cui, *Comput. Electron. Agric.* **201**, 107341 (2022). <https://doi.org/10.1016/j.compag.2022.107341>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.