

Special Issue: Artificial Intelligence

Scientific Paper

Doi: <http://dx.doi.org/10.1590/1809-4430-Eng.Agric.v42nepe20210160/2022>

## DETERMINATION OF QUALITY AND RIPENING STAGES OF 'PACOVAN' BANANAS USING VIS-NIR SPECTROSCOPY AND MACHINE LEARNING

Iara J. S. Ferreira<sup>1</sup>, Sarah L. F. de O. Almeida<sup>2</sup>,  
Acácio Figueiredo Neto<sup>2</sup>, Daniel dos Santos Costa<sup>2\*</sup>

<sup>2\*</sup>Corresponding author. College of Agricultural and Environmental Engineering, Federal University of the São Francisco Valley/ Juazeiro - BA, Brazil. E-mail: [daniel.costa@univasf.edu.br](mailto:daniel.costa@univasf.edu.br) | ORCID ID: <https://orcid.org/0000-0001-7703-3183>

### KEYWORDS

quality attributes, non-destructive method, *Musa spp.*

### ABSTRACT

This paper aimed to develop predictive models to determine total soluble solids, firmness, and ripening stages of 'Pacovan' bananas, using Vis-NIR spectroscopy and machine learning algorithms. A total of 384 bananas were divided into different days of storage (0, 3, 6, 9, 12, 15, 18, and 21 days) at two temperatures (25°C and 20°C). Bananas were subjected to spectral analysis using a spectrometer operating in spectral range of 350 – 2500 nm. Physicochemical parameters of quality, total soluble solids, and firmness were determined by reference analyses. Different machine learning algorithms were used to develop regression models and supervised classification. The best model for total soluble solids was the Random Forest with variable selection, showing an  $R^2_{cv}$  of 0.90 and RMSECV of 2.31. The best model for firmness was the Support Vector Machine with variable selection, showing an  $R^2_{cv}$  of 0.84 and RMSECV of 7.98. The best classification model for different ripening stages was the Multilayer Perceptron with variable selection, which achieved the precision of 74.22%. Therefore, Vis-NIR spectroscopy associated with machine learning algorithms is a promising tool for monitoring the quality and ripening stages of 'Pacovan' bananas.

### INTRODUCTION

Brazil is the third largest fruit producer in the world, with approximately three million hectares of planted area. In this context, bananas (*Musa spp.*) are one of the most cultivated fruits in the country. This crop is a source of fiber, vitamin C, carbohydrates, and highly important mineral nutrients, such as calcium, potassium, phosphorus, and magnesium, which contribute to human nutrition and, consequently, stimulate the commodity chain (Castilho et al., 2014; Neris et al., 2018; ABRAFRUTAS, 2019; Santos et al., 2019).

Post-harvest bananas develop climacteric processes characterized by biochemical and physical transformations, capable of directly affecting the nutritional properties and acceptability of the fruit. For that reason, during this stage, monitoring and controlling the

quality of these characteristics is important for the product commercialization (Hossain & Iqbal, 2016; Xie et al., 2018; Cho & Koseki, 2021).

Traditional methods of monitoring the post-harvest quality of bananas are usually conducted through wet methods, which are destructive, invasive, and time-consuming. These methods can properly represent the characteristics of the fruit, but they also bring food loss and waste, as well as loss of efficiency in harvesting decision-making. Thus, there is an increasing interest in non-destructive methods that provide rapid monitoring without losses (Sanaeifar et al., 2016).

Visible and near-infrared spectroscopy (Vis-NIR) is a particular case of a rapid and non-destructive technique which has been incorporated into processes to control and monitor the quality of food products. These products

<sup>1</sup> School of Agricultural Engineering, State University of Campinas/ Campinas - SP, Brazil.

<sup>2</sup> College of Agricultural and Environmental Engineering, Federal University of the São Francisco Valley/ Juazeiro - BA, Brazil.

benefit from the acquisition of relevant information concerning numerous physicochemical aspects, using optical mechanisms of interaction with the products (Hu et al., 2017).

In the fruit sector, studies have revealed the technique effectiveness to determine the quality of a wide variety of crops, such as passion fruit (Oliveira-Folador et al., 2018), grape (Costa et al., 2019), tomato (Li et al., 2021), strawberry (Xie et al., 2021), pear (Yu et al., 2021), and banana (Cho & Koseki, 2021). Different machine learning methods can be used to process and analyze data obtained by Vis-NIR spectroscopy, and then enable the quantification and classification of numberless physicochemical and colorimetric parameters.

Prediction and classification models can, for example, be processed by: inductive learning, by means of vector decision trees for characteristics of data sets (Shafiee & Minaei, 2018); parameterized learning of functions, using data separation into planes or hyperplanes in a vector space (Vapnik, 2000; Mireei et al., 2017); learning by average proximity correspondences between data (Sabanci & Akkaya, 2016); and learning by reproduction networks of different types of intelligent mechanisms (Nunes et al., 2010).

Machine learning methods, such as J48 (Quinlan, 1993), Random Forest, Support Vector Machine (Keerthi et al., 2001), K-Nearest Neighbours, and Artificial Neural Network are commonly used for developing prediction and classification models. However, their use for spectroscopy data in agricultural and food products is still little explored. Hence, the objective of this paper is to develop predictive

models to determine total soluble solids, firmness, and ripening stages of 'Pacovan' bananas, using Vis-NIR spectroscopy associated with machine learning algorithms.

## MATERIAL AND METHODS

A total of 384 'Pacovan' cultivar bananas at the green ripening stage were obtained at Central de Abastecimento Agrícola, in the municipality of Juazeiro (Juazeiro, Bahia, Brazil: Latitude: 09° 24' 42" S and Longitude: 40° 29' 55" W). They were subjected to surface sterilization and stored at temperatures ranging from 20 to 25°C, and relative humidity between 50 and 58%. The bananas were evaluated on different days of storage (0, 3, 6, 9, 12, 15, 18, and 21 days). On each evaluation day, 48 bananas were individually subjected to a process of spectral acquisition, followed by reference analyses for total soluble solids (TSS) and firmness.

The spectral acquisition system consisted of: (1) a FieldSpec 3 spectrometer (Analytical Spectral Devices, Boulder, Colorado, USA), with an 8° optical sensor field of view, wavelength range from 350 to 2500 nm, resolution from 3 to 10 nm, and precision of  $\pm 1$  nm; (2) 50-W quartz-tungsten-halogen light source; (3) a dark chamber with dimensions of 100 x 50 x 50 cm; and (4) a computer with RS3 software (Analytical Spectral Devices, Boulder, Colorado, USA) (Figure 1). A white Spectralon ceramic plaque (Labsphere Inc., North Sutton, NH, USA) with approximately 100% reflectance was used as a calibration standard, and each spectrum was established by the average of sixty scans performed on two sides of the banana.

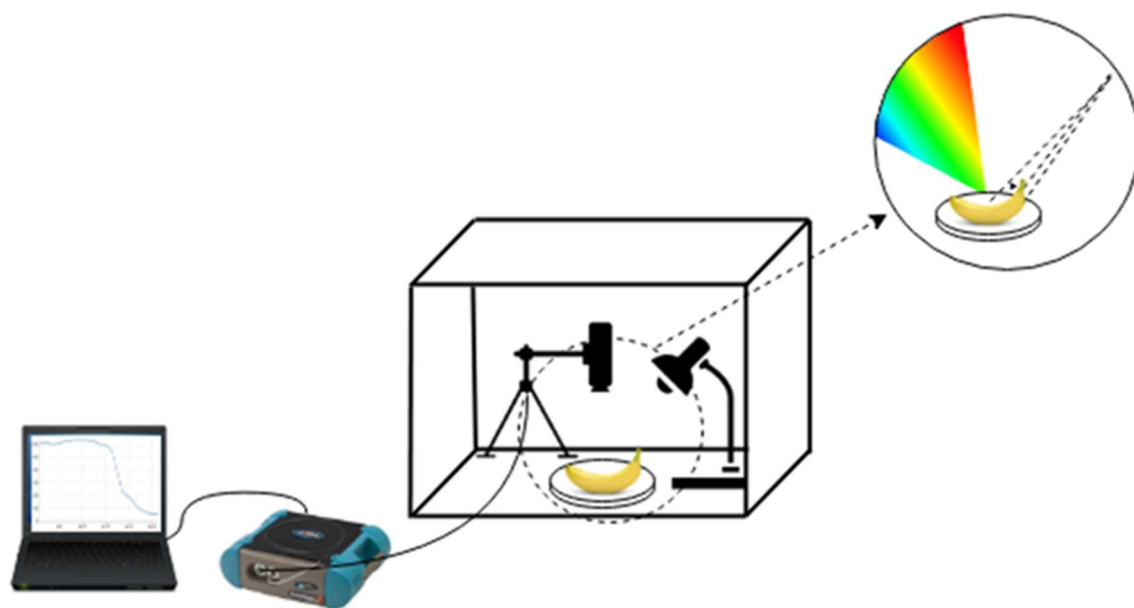


FIGURE 1. Vis-NIR reflectance spectra acquisition system, consisting of a spectrometer, light source, dark chamber, and computer.

On each evaluation day, visual classification of fruits at the green, nearly ripe, ripe, and rotting ripening stages was carried out based on and adapted from the Von Loesecke (1950) ripening scale. Total soluble solids (TSS) were determined from the fruit juice, using a digital refractometer (HI 96804, Hanna Instruments, USA) with a measuring range between 0 and 85%, precision of  $\pm 0.2\%$ , and the results were expressed as percentages (%). Firmness was measured by a digital penetrometer (PTR-300, Instrutherm, Brazil) with a 5 mm cylindrical probe,

and the results were expressed as Newtons (N). Refractometer and penetrometer readings were performed only once on each evaluated fruit.

Predictive models were developed with classifiers J48, Random Forest (RF), Support Vector Machine using Sequential Minimal Optimization (SVM-SMO), K-Nearest Neighbours algorithm (KNN-IBK), and Multilayer Perceptron Artificial Neural Network (MLP-ANN), within the Weka 3.8.4 software implementation environment (University of Waikato, New Zealand). Models were

developed with default settings of the algorithms in the software and based on 10-fold and leave-one-out cross-validation methods.

To eliminate possible collinearity problems within the spectral data set, a variable selection was conducted using the following filters: Correlation-based Feature Selection (CFS) (Hall, 1999) with the Best-first search method and CfsSubsetEval function; and Wrapper (Kohavi & John, 1997) with the GreedyStepwise search method and WrapperSubsetEval function.

Using the CFS filter, the selection was carried out investigating characteristics between variables by correlation functions, whereas the selection with the Wrapper filter was performed using modeling algorithms to test variable sets (Witten et al., 2011). Models were reconstructed with the selected variables. Predictor variables were the full-spectrum data and selected variables, while response variables were TSS, firmness, and ripening stages.

Regression models performance was evaluated through the coefficient of determination of cross-validation ( $R^2_{cv}$ ) (Equation 1), root mean square error of cross-validation (RMSECV) (Equation 2), and mean absolute error of cross-validation (MAECV) (Equation 3). In addition, the performance of supervised classification models was evaluated using precision (P) (Equation 4), Kappa concordance index (K) (Equation 5), sensitivity (SEN) (Equation 6), selectivity (S) (Equation 7), and false-positive rates (FPR) (Equation 8).

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

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

$$MAECV = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n} \quad (3)$$

$$P = \frac{TP + TN}{TP + FN + FP + TN} * 100 \quad (4)$$

$$K = \frac{Po - Pc}{1 - Pc} \quad (5)$$

$$SEN = \frac{TP}{TP + FN} * 100 \quad (6)$$

$$S = \frac{TN}{FP + TN} * 100 \quad (7)$$

$$FPR = \frac{FP}{FP + TN} * 100 \quad (8)$$

Where:

$\hat{y}_i$  - Values predicted by the cross-validation model;

$y_i$  - Reference values;

$\bar{y}$  - Mean of the predicted values;

n - Number of cross-validation samples;

$\sigma_r$  - Standard deviation of reference values;

$\sigma_p$  - Standard deviation of predicted values;

TP - True positives;

TN - True negatives;

FP - False positives;

FN - False negatives;

Po - Overall accuracy,

Pc - Proportion of units for which agreement is expected by chance.

The Area Under the Curve (AUC) parameter, created during model processing using the graphical technique of visualization, organization, and selection of Receiver Operating Characteristic (ROC) classifiers (Prati et al., 2008), was used to indicate the discriminative capacity of the models. An AUC value close to 1 indicates high discriminative capacity, while a value close to 0.5 means little discriminative power (Luo et al., 2012).

## RESULTS AND DISCUSSION

Total soluble solids content increased from 1.8% to 35.5%, while firmness decreased from 86.6 N to 3.3 N over the days of storage at any temperature (Figure 2). Biochemical and cellular alterations occur during the fruit ripening process, such as the conversion of starch into sugars, and softening of cell wall structures through the breakdown of pectins by enzymes (Yang et al., 2019; Cho & Koseki, 2021).

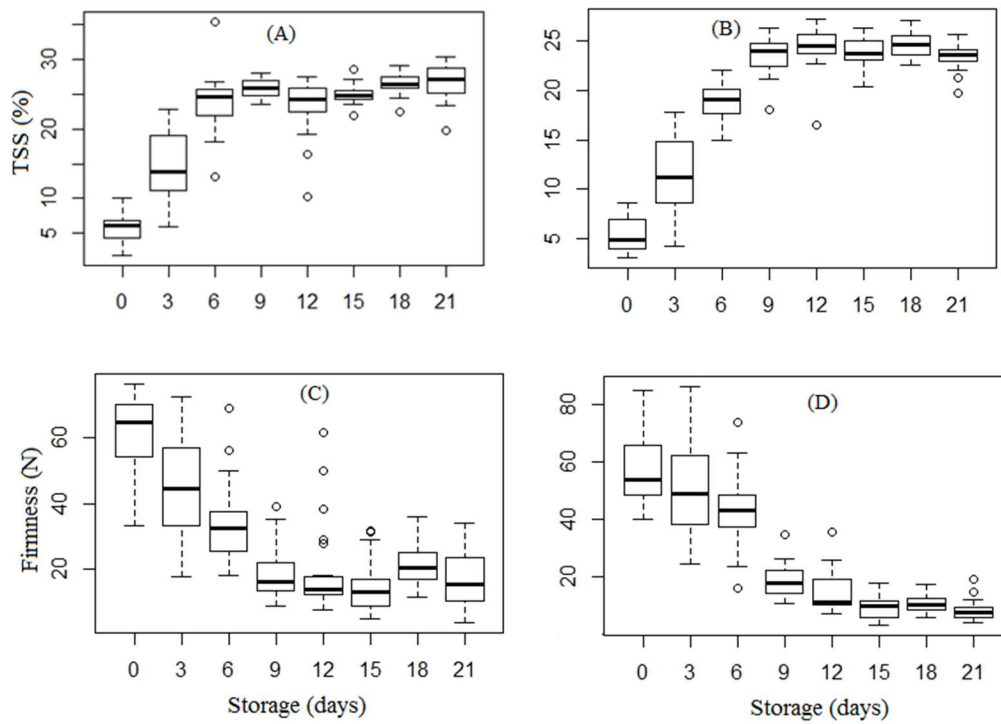


FIGURE 2. Behavior of total soluble solids and firmness of 'Pacovan' bananas during storage at 25°C (A and C) and 20°C (B and D) over the days of storage.

Figure 3 shows the behavior of the absorption spectra in the region from 400 to 2400 nm over the days of storage. In the visible region between 560 and 720 nm absorption peaks are observed for days 0 and 3 of storage. During this period, the fruit was in the initial ripening stage, and wavelengths between 600 and 750 nm account for the presence of chlorophyll, which absorbs red light (Liu et al., 2008). Subsequently, colorimetric changes occur, such as chlorophyll degradation and progressive increase in pigment compounds (carotenoids and anthocyanins) in the epidermis, as well as biochemical

changes, such as the hydrolysis of starch and cell wall components (pectins and hemicelluloses). Consequently, the banana color changed from green to yellow, and red light absorption decreased (Liu et al., 2008; Quevedo et al., 2008; Carvalho et al., 2011; Liew & Lau, 2012; Hailu et al., 2013; Adebayo et al., 2016). Wavelengths of 970, 1180, and 1440 nm showed absorption peaks on all days of storage. These wavelengths in the near-infrared region are related to the absorption bands by sugars (980 nm - second OH overtone) and water (973, 1324, and 1581 nm - third OH overtone) (Shafiee & Minaei, 2018; Costa et al., 2019).

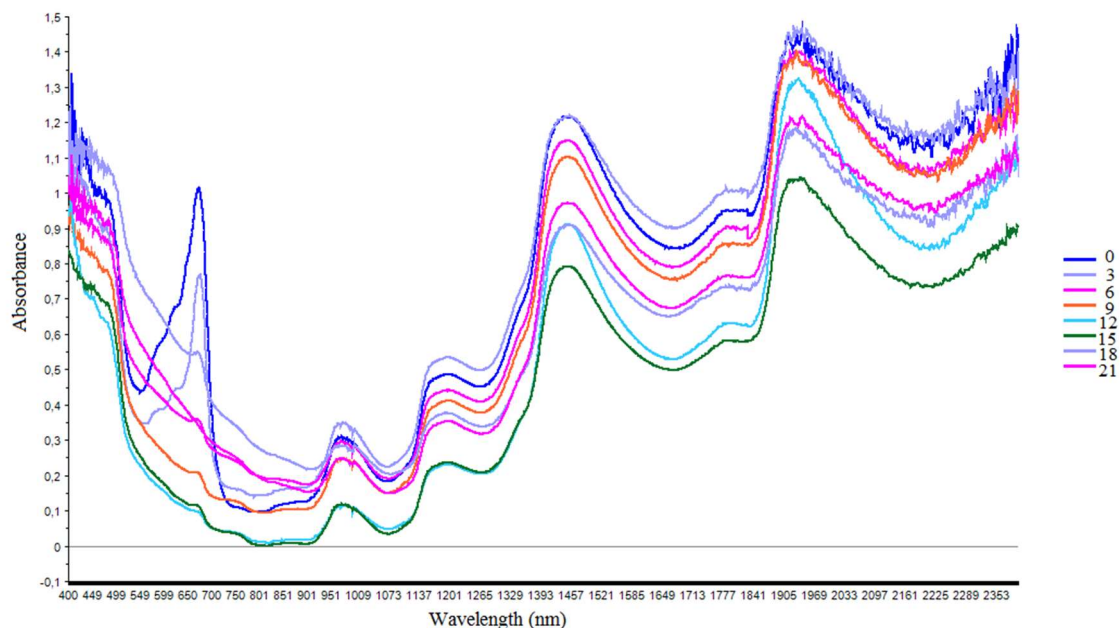


FIGURE 3. Absorbance spectra of 'Pacovan' bananas over the days of storage. Curves under different colors correspond to the days of storage.

The RF and SVM models, associated with the Wrapper selection filter, had performance for total soluble solids and firmness of 0.90 and 0.84 for  $R^2_{cv}$ , 2.31 and 7.98 for RMSECV, and 1.71 and 6.12 for MAECV, respectively (Table 1). The results of this study are similar

to those of Liew & Lau (2012), who developed models for total soluble solids and firmness of Cavendish bananas with  $R^2$  of 0.96 and 0.86, respectively, and Jaiswal et al. (2012), who developed models for total soluble solids of Grand Naine bananas with  $R^2$  of 0.88.

TABLE 1. Regression models for total soluble solids and firmness of 'Pacovan' bananas using full spectrum and variable selection.

Quality parameter	Selection filter	Model											
		RF			SVM			KNN			MLP		
		$R^2_{cv}$	RMSECV	MAECV	$R^2_{cv}$	RMSECV	MAECV	$R^2_{cv}$	RMSECV	MAECV	$R^2_{cv}$	RMSECV	MAECV
TSS	no selection	0.79	3.37	2.41	0.85	2.85	2.13	0.79	3.40	2.35	0.61	4.78	3.24
	Wrapper	<b>0.90</b>	<b>2.31</b>	<b>1.71</b>	0.85	2.93	2.32	0.88	2.48	1.85	0.83	3.26	2.46
	CFS	0.86	2.71	2.00	0.69	4.13	3.30	0.86	2.63	1.96	0.81	3.38	2.52
Firmness	no selection	0.69	11.00	8.65	0.74	10.27	7.56	0.71	10.68	8.31	0.62	12.72	9.58
	Wrapper	0.82	8.52	6.50	<b>0.84</b>	<b>7.98</b>	<b>6.12</b>	0.82	8.57	6.64	0.74	10.44	8.33
	CFS	0.75	9.86	7.55	0.70	10.88	8.54	0.78	9.39	7.39	0.66	11.70	9.24

Legend: RF – Random Forest; SVM – Support Vector Machine; KNN – K-Nearest Neighbours; MLP – Multilayer Perceptron;  $R^2_{cv}$  – coefficient of determination of cross-validation; RMSECV – root mean square error of cross-validation; MAECV – mean absolute error of cross-validation.

The Wrapper filter selected the wavelengths of 356, 404, 462, 660, 671, 727, 755, and 2196 nm for total soluble solids, and 357, 363, 365, 371, 373, 374, 378, 392, 394, 400, 402, 413, 420, 425, 457, 463, 465, 468, 469, 475, 489, 615, 646, 654, 676, 694, 2439, and 2441 nm for firmness. The large number of wavelengths selected in the visible region indicates that, for 'Pacovan' bananas, variations in starch content, which result in an expansion of soluble solids, and pectic changes in the fruit texture, which result in firmness alterations, were directly associated with colorimetric characteristics in the ripening process (Barbosa et al., 2019). The increase in performance obtained through models with selected

wavelengths becomes an advantage to determine total soluble solids and firmness, since it minimizes the complexity and computational cost of processing (Wu et al., 2010; Munera et al., 2017).

The performance of supervised classification models is indicated in Table 2. The MLP model, associated with the Wrapper filter, discriminated the different ripening stages with a precision of 74.22%, Kappa index of 0.62, sensitivity of 78.26%, selectivity of 89.31%, and false-positive rate of 10.69%. The precision is close to that reported by Mustafa et al. (2009), who developed models with a precision of 81% to discriminate ripening stages of bananas.

TABLE 2. Classification models for ripening stages of 'Pacovan' bananas using full spectrum and variable selection.

Model	Selection filter	precision (%)	Kappa	sensitivity (%)	selectivity (%)	FPR (%)
J48	no selection	62.50	0.42	60.52	76.86	23.14
	Wrapper	71.09	0.57	77.01	87.96	12.04
	CFS	71.87	0.59	76.43	88.25	11.75
RF	no selection	65.36	0.49	63.49	84.46	15.54
	Wrapper	68.75	0.51	81.02	88.04	11.96
	CFS	72.39	0.59	77.12	88.44	11.56
SVM	no selection	72.13	0.59	76.20	88.27	11.73
	Wrapper	73.18	0.60	78.53	89.25	10.75
	CFS	63.54	0.44	70.47	86.09	13.91
KNN	no selection	64.58	0.47	59.15	84.18	15.82
	Wrapper	73.96	0.62	76.22	89.02	10.98
	CFS	73.44	0.61	77.65	88.94	11.06
MLP	no selection	44.01	0.20	42.98	70.63	29.37
	Wrapper	<b>74.22</b>	<b>0.62</b>	<b>78.26</b>	<b>89.31</b>	<b>10.69</b>
	CFS	70.83	0.57	76.86	87.82	12.18

Legend: RF – Random Forest; SVM – Support Vector Machine; KNN – K-Nearest Neighbours; MLP – Multilayer Perceptron; FPR – false-positive rates.

The Wrapper filter selected the wavelengths of 357, 364, 432, and 934 nm for the MLP classifier. The selection of wavelengths improved the model discrimination capacity, reducing errors caused by temporal variabilities, and removed redundant information that compromises the classifier performance (Ramos, 2003).

The classifier was able to discriminate the green ripening stage with 100% accuracy, and had a satisfactory performance in discriminating the other stages (Table 3).

TABLE 3. Precision (%) and Area Under the Curve (%) performance of the MLP model for individual discrimination of ripening stages of 'Pacovan' bananas.

MLP model	Green	Nearly ripe	Ripe	Rotting
Precision	100.00	73.50	70.70	68.90
ROC area	100.00	89.90	82.80	83.00

Legend: MLP – Multilayer Perceptron.

## CONCLUSIONS

Regression models based on the Random Forest and Support Vector Machine algorithms with variable selection were able to predict total soluble solids and firmness with performances around 90%. The supervised classification model based on the Multilayer Perceptron algorithm with variable selection discriminated the different ripening stages of the bananas with a precision greater than 70%. The selection of wavelengths in the visible and near-infrared spectral regions significantly increased the models capacity to determine the total soluble solids, firmness, and ripening stages. Therefore, Vis-NIR spectroscopy associated with machine learning algorithms is a promising tool for monitoring the quality attributes and ripening stages of 'Pacovan' bananas.

## REFERENCES

ABRAFRUTAS - Associação Brasileira dos Produtores Exportadores de Frutas e Derivados (2019) Os rumos da produção de frutas no Brasil.

Adebayo SE, Hashim N, Abdan K, Hanafi M, Mollazade K (2016) Prediction of quality attributes and ripeness classification of bananas using optical properties. *Scientia Horticulturae* 212:171-182.

Barbosa LFS, Alves AL, Sousa KSM, Neto AF, Cavalcante IHL, Vieira JF (2019) Qualidade pós-colheita de banana 'Pacovan' sob diferentes condições de Armazenamento. *Magistra* 30:28 – 36.

Carvalho AV, Seccadio LL, Jr MM, Nascimento WMO (2011) Qualidade pós-colheita de cultivares de bananeira do grupo 'maçã', na região de Belém – PA. *Revista Brasileira de Fruticultura* 33:1095-1102.

Castilho LG, Alcantara BM, Clemente E (2014) Desenvolvimento e análise físico-química da farinha da casca, da casca in natura e da polpa de Banana verde das cultivares maçã e prata. *E-xacta* 7:107-114.

Cho B, Koseki S (2021) Determination of banana quality indices during the ripening process at different temperatures using smartphone images and an artificial neural network. *Scientia Horticulturae* 288:110382.

Adebayo et al. (2016) state that there is a negative correlation between ripening stages and absorption coefficients of photons from the light falling on the banana. As the fruit goes through the ripening stages, absorption rates of these photons, which are related to chemical compositions, such as sugar content, soluble solids content, and the presence of photosynthesis compounds, are reduced (Adebayo et al., 2016).

Costa DS, Mesa NFO, Freire MS, Ramos RP, Mederos BJT (2019) Development of predictive models for quality and maturation stage attributes of wine grapes using VIS-NIR reflectance spectroscopy. *Postharvest Biology and Technology* 150:166–178.

Hailu M, Workneh TS, Belew D (2013) Review on postharvest technology of banana fruit. *African Journal of Biotechnology* 12:635-647.

Hall MA (1999) Correlation-based feature subset selection for machine learning. Hamilton, New Zealand.

Hossain, M S, Iqbal A (2016) Effect of shrimp chitosan coating on postharvest quality of banana (*Musa sapientum* L.) fruits. *International Food Research Journal* 23: 277-283.

Hu J, Ma X, Liu L, Wu Y, Ouyang J (2017) Rapid evaluation of the quality of chestnuts using near-infrared reflectance spectroscopy. *Food Chemistry* 231:141–147.

Jaiswal P, Jha SN, Bharadwaj R (2012) Non-destructive prediction of quality of intact banana using spectroscopy. *Scientia Horticulturae* 135:14–22.

Keerthi SS, Shevade SK, Bhattacharyya C, Murthy KRK (2001) Improvements to Platt's SMO algorithm for SVM classifier design. *Neural Computing* 13:637-649.

Kohavi R, John GH (1997) Wrappers for feature subset selection. *Artificial Intelligence* 97:273-324.

Li X, Tsuta M, Hayakawa F, Nakano Y, Kazami Y, Ikehata Akifumi (2021) Estimating the sensory qualities of tomatoes using visible and near-infrared spectroscopy and interpretation based on gas chromatography-mass spectrometry metabolomics. *Food Chemistry* 343:128470.

Liew CY, Lau CY (2012) Determination of quality parameters in Cavendish banana during ripening by NIR spectroscopy. *International Food Research Journal* 19:751/758.

Liu Y, Chen W, Ouyang A (2008) Nondestructive determination of pear internal quality indices by visible and near-infrared spectrometry. *Food Science and Technology* 41(9): 1720-1725.

- Luo X, Takahashi T, Kyo K, Zhang S (2012) Wavelength selection in vis/NIR spectra for detection of bruises on apples by ROC analysis *Journal of Food Engineering* 109:457–466.
- Mireei SA, Amini-Pozveh S, Nazeri M (2017) Selecting optimal wavelengths for detection of insect infested tomatoes based on SIMCA-aided CFS algorithm. *Postharvest Biology and Technology* 123:22–32.
- Munera, S, Amigo JM, Blasco J, Cubero S, Talens P, Aleixos N (2017) Ripeness monitoring of two cultivars of nectarine using VIS-NIR hyperspectral reflectance imaging. *Journal of Food Engineering* 214:29–39.
- Mustafa NBA, Ahmed SK, Ali Z, Yit WB, Abidin AAZ, Sharrif ZAM (2009) Agricultural Produce Sorting and Grading using Support Vector Machines and Fuzzy Logic. *International Conference on Signal and Image Processing Applications* 391-396.
- Neris TS, Silva SS, Loss RA, Carvalho JWP, Guedes SF (2018) Avaliação físico-química da casca da banana (*Musa spp.*) in natura e desidratada em diferentes estádios de maturação. *Ciência e Sustentabilidade* 4:5-21.
- Nunes I, Spatti DH, Flauzino RA (2010) Redes neurais artificiais para engenharia e ciências aplicadas. São Paulo, artliber. Editora artliber, p399.
- Oliveira-Folador G, Bicudo MO, Andrade EF, Renard CMGC, Bureau S, Castilhos F (2018) Quality traits prediction of the passion fruit pulp using NIR and MIR spectroscopy. *Food Science and Technology* 95:172-178.
- Prati RC, Batista G, Monard MC (2008) Curvas ROC para avaliação de classificadores. *IEEE America Latina* 6:1-8.
- Quevedo R, Mendoza F, Aguilera JM, Chanona J, Gutiérrez-López G (2008) Determination of senescent spotting in banana (*Musa cavendish*) using fractal texture Fourier image. *Journal Food Engineering* 84:509–515.
- Quinlan R (1993) C4.5: Programs for machine learning. San Mateo, Morgan Kaufmann Publishers, 16:235-240.
- Ramos JP (2003) Redes neurais artificiais na classificação de frutos: cenário bidimensional. *Ciência e agrotecnologia* 27 (2):356-362.
- Sabanci K, Akkaya M (2016) Classification of Different Wheat Varieties by Using Data Mining Algorithms. *International Journal of Intelligent Systems and Applications in Engineering* 4:40–44.
- Sanaeifar A, Bakhshipour A, de la Guardia M (2016) Prediction of banana quality indices from color features using support vector regression. *Talanta* 148:54–61.
- Santos WWV, Silva KRO, Barbosa RC, Oliveira JB, Silva JSA, Medeiros EV (2019) Efeito de diferentes métodos de maturação sobre a qualidade da banana prata. *Diversitas Journal* 4:1092-1104.
- Shafiee S, Minaei S (2018) Combined data mining/NIR spectroscopy for purity assessment of lime juice. *Infrared Physics & Technology* 91:193–199.
- Vapnik, V (2000) *The Nature of statistical learning theory*. New York, Springer-Verlag.
- Von Loesecke H W (1950) *Bananas*. New York, Interscience Publishers, p52-66.
- Xie C, Chu B, He Y (2018) Prediction of banana color and firmness using a novel wavelengths selection method of hyperspectral imaging. *Food Chemistry* 245:132–140.
- Xie D, Liu D, Guo W (2021) Relationship of the optical properties with soluble solids content and moisture content of strawberry during ripening. *Postharvest Biology and Technology* 179:111569.
- Witten IH, Frank E, Hall MA (2011) *Data mining: practical machine learning tools and techniques*. San Francisco, Morgan Kaufmann.
- Wu D, He Y, Nie P, Cao F, Bao Y (2010) Hybrid variable selection in visible and near-infrared spectral analysis for non-invasive quality determination of grape juice. *Analytica Chimica Acta* 659:229–237.
- Yang J, Zeng J, Wen L, Zhu H, Jiang Y, John A, Yu L, Yang B (2019) Effect of morin on the degradation of water-soluble polysaccharides in banana during softening. *Food Chemistry* 287:346–353.