Journal of Agriculture and Food Research 14 (2023) 100779

Contents lists available at ScienceDirect

Journal of Agriculture and Food Research

journal homepage: www.sciencedirect.com/journal/journal-of-agriculture-and-food-research

Detection of hollow heart disorder in watermelons using vibrational test and machine learning

F.J. Simón-Portillo, D. Abellán-López *, M. Fabra-Rodriguez, R. Peral-Orts, M. Sánchez-Lozano

Department of Mechanical and Energy Engineering, Miguel Hernandez University of Elche, Avda. de la Universidad, Elche, 03202, Spain

ARTICLE INFO

Keywords:
Watermelon
Non-destructive testing
Vibrational method
Hollow detection
Classifier algorithms
Machine learning

ABSTRACT

The presence of internal voids in watermelons has an impact on the costs of producers and on consumer confidence. Various studies have shown that the vibrational parameters of the fruit are related to maturity, quality and the existence of internal defects. A method for the detection of internal voids in seedless watermelons based on vibrational parameters obtained in impact hammer tests and machine learning is presented. After a statistical study of the test results, the frequency of the first peak of the vibrational response and the density of the watermelon are selected as predictors to be used in the classification algorithms. The accuracy of detecting hollow watermelons increases if firmness estimator is introduced as a predictor. Probabilities of success above 89% in the detection of internal voids have been achieved using different classification algorithm.

1. Introduction

The presence of voids in watermelons affects internal texture and taste. These defective watermelons are returned and not charged for, resulting in costs to the producer, complaints and loss of customer confidence. The voids inside the watermelon, a defect known as *hollow heart*, in some cases reach 50% of the internal volume of the watermelon and can be caused by irregular growth between the centre and the outside. This can occur when atmospheric conditions during growth alternate from wet to dry or when there are large temperature changes [1]. Overwatering or excessive nitrogen fertilisation can also cause this type of defect [2,3].

The marketing of hollow watermelons entails an additional cost for producers. Watermelons cannot be inspected invasively so multiple research projects have been carried out in order to find a non-destructive inspection method [4,5]. Some of these methods are based on vibrational and/or acoustic techniques. On the packing line, or even in the field, there tend to be trained experts responsible for the initial screening of watermelons with internal defects. This method consists of manually tapping each watermelon so that an expert operator can classify it based on the sound produced. Although deeper sounds are associated with the existence of voids, the method is still subjective and imprecise, as well as involving a high cost in terms of time and personnel.

Automating a method for detecting internal defects based on vibrational techniques presents certain difficulties. Firstly, the frequency and magnitude with which an object responds to vibrational excitation depends on its elasticity, density, size and shape. These properties are highly variable within the same batch of watermelons and depend on factors such as maturation time, water level, rind thickness, etc. Secondly, watermelons of the same variety grown in the same area and season may present internal voids of very different geometries and sizes [1]. Fig. 1 shows various watermelons used in this study with different types and degrees of internal voids. How these voids affect the vibrational characteristics of the watermelons as a function of their shape and size is a field that has yet to be fully studied.

The correlation between the natural frequencies of watermelons obtained by acoustic impulse and their firmness level was studied for first time by Yamamoto et al. [6]. Yamamoto found that firmness can be estimated with indices based on frequency of resonance and watermelon mass. These expressions are modifications of those used previously by other authors to determine the state of other types of fruit [7,8]. In Yamamoto study, watermelons were excited by tapping them with a wooden pendulum and recorded the sound produced. They found no appreciable differences when the angle of the pendulum was varied. Yamamoto also found that the storage time of the watermelons affected their natural frequencies, phenomenon that has been studied later by other authors [9]. Other authors have carried out similar studies to determine the firmness and maturity of watermelons using non-destructive acoustic methods [10–13]. Some of these studies use the relationship between the mass properties and natural frequencies of

* Corresponding author.
E-mail address: dabellan@umh.es (D. Abellán-López).

https://doi.org/10.1016/j.jafr.2023.100779
Received 10 May 2023; Received in revised form 14 August 2023; Accepted 5 September 2023
Available online 7 September 2023
2666-1543/

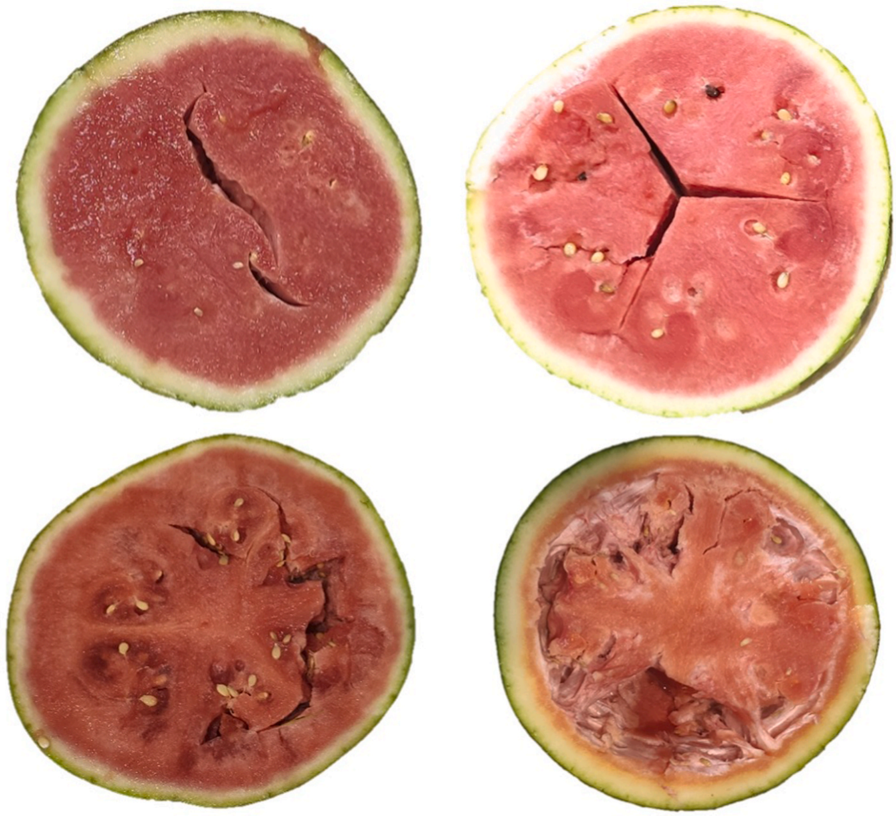

Fig. 1. Examples of watermelons with internal defects used in this study.

watermelons to predict the ripeness level [3,14,15], corroborating Yamamoto's findings in the 1980s. He et al. [16] conducted a vibrational study hitting watermelons with a pendulum and they found correlation between natural frequencies and the soluble solids content (SSC).

Other studies have attempted to determine the maturity level of watermelons based on vibrational parameters other than natural frequencies, for example, sound transmission velocity [17], or on the attenuation coefficient of an acoustic wave [18]. Mizrach et al. [18] also found a correlation between the attenuation coefficient and the soluble solids content (SSC) of watermelons.

Various papers have studied the relationship between the vibrational parameters of watermelons and the presence of internal voids [3,15,19]. In the study carried out for Stone et al. [3] subjected the fruit, placed on the floor, to an acoustic pulse and measured the magnitude of the response of the fruit. They found a correlation between the magnitude of the proposed frequency bands and the Melon Hollow Heart Index. They obtained better results when the test was performed on the floor than when the melon was placed on a solid base. These authors found a correlation between acoustic impedance and the level of firmness and sugar content. They also found a correlation between the energy contained in the 80–130 Hz range and the existence of voids. Stone et al. [3] found that the results of internal void prediction improve if firmness values measured with experimental techniques are introduced as a parameter.

Iglesias et al. [20] attempted to detect the existence of internal voids in watermelons by recording the response of a melon subjected to a light impact with a microphone. In 2006, a method was developed by Noh and Choi that allowed for the detection of internal voids with an accuracy of 90% in laboratory. The method utilizes an acoustic impulse and measures vibration using an optical sensor [21].

Apart from methods based on vibro-acoustic techniques, there are other non-destructive methods for detecting internal defects in watermelons and other fruits for which other types of technologies are used. These techniques include X-ray, tomography or spectroscopy [4,5,22]. An automatic system to classify watermelons based on their density was developed by Kato [23]. Kato successfully determined the absence of voids in a sample of 59 defect-free and 16 hollow watermelons. He found that density in defect-free watermelons is also related to soluble solids content.

Vibrational methods have also been used to determine the state of maturation in other types of fruit than watermelons and other papers study alternative parameters such as the magnitude of the frequency band, damping coefficient or the pulse propagation velocity [24,25].

Increasingly, classification techniques based on machine learning are being developed, which include k-Nearest Neighbour (KNN), neural network, neural network, Support Vector Machine (SVM), and the decision tree [26]. Multiple techniques can be used to classify watermelons according to their ripeness level based on vibro-acoustic test data. Abbaszadeh et al. [27] use a KNN classifier and vibrational test data to detect watermelon ripeness level with a 95% success rate. Chawgien and Kiattisin [28] test a large number of classifier algorithms to determine the maturity level. Alipasandi et al. [29] combine vibro-acoustic testing with KNN and SVM algorithms. Machine learning technology has also been used with other types of watermelon tests, for example, Liu et al. [30] use several algorithms (random forest, neural network, SVM, principal components) to measure and classify watermelons according to their quality by multispectral imaging, and Arboleda et al. [31] use infrared spectroscopy and SVM classifier to determine the ripeness of watermelons. However, the use of classifier algorithms for internal voids detection has not been studied in previous works.

In the study described in this document, an analysis of the relationship between the existence or not of voids and the physical properties of watermelons and the vibrational parameters obtained by vibratory tests with an impact hammer, was performed. The objective was to determine the best parameters to be used as predictors of the existence of voids by

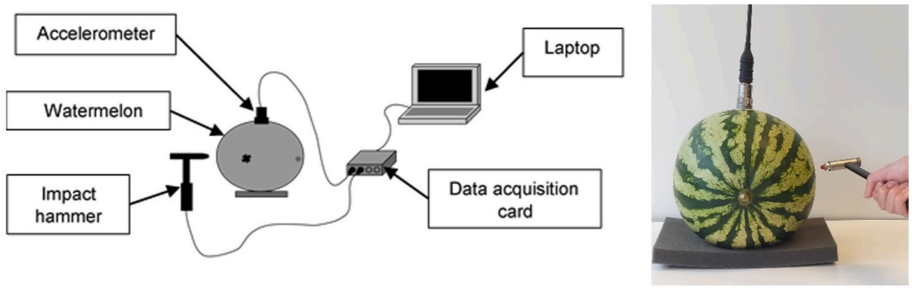

Fig. 2. Configuration of the frequency test on watermelons. Left: schematic. Right: image of the test.

means of statistical analysis. Subsequently, classification algorithms were used to evaluate the effectiveness of these parameters for classifying watermelons as a function of the presence of internal defects. The combined application of vibrational techniques and machine learning for the detection of voids in watermelons represents a novelty that can help enhance accuracy in the detection these internal defects.

2. Materials and methods

2.1. Vibrational tests

For the statistical study, 120 seedless watermelons harvested between May and June 2022 were used. These watermelons were used to train the classifier algorithms. The company that supplied the watermelons screens the watermelons when they arrive at the warehouse by means of operators specialised in detecting internal voids manually. Part of the sample for this study was taken from the group of watermelons that had previously been classified as potentially hollow. Another part of the sample was taken from the group of watermelons classified as good quality by the agricultural company AGRICOLA NAVARRO DE HARO S. L [32]. For this reason, the sample used contains a higher proportion of hollow watermelons than usually detected during the harvest. This was done in order to have a sufficiently large number of watermelons with internal defects for the statistical analysis, see section 2.3, without the need to test thousands of watermelons. Of the 120 watermelons used for classifier training, 25 were from the hollow watermelons discarded by the experts.

Prior to testing, the watermelons were stored in a conditioned room at 25°C for a minimum of 24 h. Before the vibrational test, each watermelon was weighed with a dynamometer AMETEK CHATILLON DFX-050 and its volume was estimated. To determine watermelon volume, perimeters were measured with a tape measure in the equatorial plane and in a plane perpendicular to it, passing through the pedicle. For volume estimation, the watermelon shape was approximated to the ellipsoid of revolution.

The configuration for Fig. 2 was used for the frequency test. Each watermelon was placed with the pedicle in a horizontal position on a 25 mm polyurethane base with a density of 30 kg m^{-3}. A uniaxial accelerometer, model PCB 601A01, connected to the first channel of the NI 9233 data acquisition card was attached to the top of the watermelon, which is previously cleaned with alcohol. The watermelon was tapped at the equator, at the height of the pedicle, with a PCB 086C03 impact hammer connected to the second channel of the acquisition card. The impact hammer was fitted with a soft rubber tip in order not to damage the watermelon skin during the test. The acquisition card was connected to a laptop computer for data recording, pre-processing and storage using LabVIEW software.

The acquisition card recorded signals of 1 s duration from the impact of the hammer with a sampling frequency of 2 kHz. The accelerometer signal y(t) was subjected to Fast Fourier Transform (FFT) to obtain its frequency content Y(ω). The impact hammer has a load cell at its tip which records the applied force signal x(t). In order to perform a modal analysis, all frequencies of the test specimen must be excited equally, which is not the case with the method used as an infinitely small impact duration would be required. The effect of the actual impact duration can be corrected by transforming the input signal X(ω) and thus calculating the frequency response function H(ω) according to equation (1) [33].

$$H(\omega) = \frac{Y(\omega)}{X(\omega)} \tag{1}$$

The data recorded by both the hammer and the accelerometer as well as the frequency response were exported to Matlab for subsequent analysis. Each watermelon was tested 5 times using the same procedure

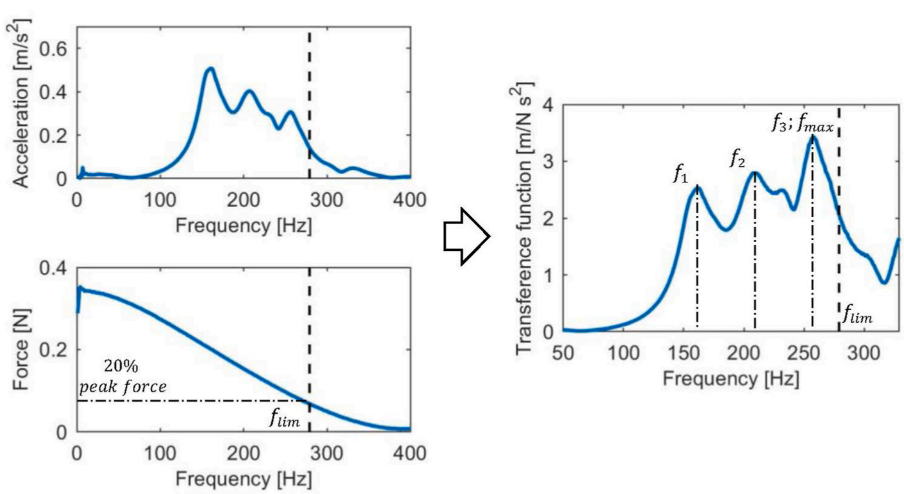

Fig. 3. Watermelon vibrational test signal processing.

Table 1
Correlation coefficients between the variables studied.

	Mass	Apparent density	Radius (a)	f_{max}	f_1	$firmness_a$	$firmness_c$
Mass	1.00	−0.10	0.95	−0.43	−0.49	0.45	0.14
Apparent density	–	1.00	−0.27	0.45	0.46	0.37	0.44
Radius (a)	–	–	1.00	−0.50	−0.57	0.32	0.02
f_{max}	–	–	–	1.00	0.69	0.30	0.48
f_1	–	–	–	–	1.00	0.55	0.79
$firmness_a$	–	–	–	–	–	1.00	0.95
$firmness_c$	–	–	–	–	–	–	1.00

in order to detect any experimental errors. Finally, the real state of each watermelon pulp was checked by opening it, and all of them were classified into watermelons with internal defects or watermelons without internal defects.

2.2. Data processing

Subsequently, the accelerometer and hammer signals were processed in the frequency domain using Matlab software in order to obtain the vibrational parameters that best predict the internal state of the watermelons. Fig. 3 shows the typical curve of the force signal captured by the impact hammer in the frequency domain. According to expression (*1*), the acceleration pulse values for high frequencies are over-magnified due to the fact that the convolution is performed with excitation values that are too low. Therefore, only the data up to the frequency at which the hammer signal falls below 20% of its maximum value, which is called $f_{\lim}$, have been used in the analysis.

Fig. 3 also shows the spectral response function curve for a watermelon tested as described in the previous section. The peaks in the acceleration signal at frequencies close to zero can be discarded, as they are associated with free solid movements and not with the transmission of vibrations inside the fruit. This assumption is valid if the frequencies of the discarded peaks are one order of magnitude lower than the frequencies studied, which is true for the acceleration signals in this study. Therefore, the signal range used for the analysis was between 50 Hz and the limiting frequency $f_{\lim}$. In this frequency range, the frequencies at which the first three peaks in the transfer function occur had been obtained, named f_1, f_2 and f_3 respectively. In some cases, watermelons only exhibited two peaks in the frequency range studied so only these first two frequencies were used. It had been found that the frequencies at which the peaks occur are not influenced by the point of impact, provided that it was made on the equator of the watermelon in the direction approximately perpendicular to the direction of measurement of the uniaxial accelerometer. Of these 3 frequencies, the frequency at which the highest amplitude peak in the transfer function has been named f_{max}.

The results of all samples were analysed, discarding those watermelons for which anomalous data were obtained in the accelerometer or hammer signals. The most common error was due to the accelerometer lift-off. In those cases, the shape of some signals didn't correspond to the typical curve of an acceleration spectrum shown in Fig. 3. Anomalous curves had a huge ripple or no peaks at any frequency within the analysed range. If after discarding the questionable signals there were not at least three acceleration curves, the watermelon was discarded for the analysis. Finally, the characteristic frequency values for each watermelon were calculated as the median of those obtained from the repeated watermelon tests.

After screening out watermelons that showed dubious data, results were available for 117 watermelons, 22 of which were found to be hollow and 95 without internal defects.

The maturation level of fruit has been demonstrated to affect their natural frequencies by several authors. Abbot et al. [8] used vibrational analysis to estimate the level of maturation in both fruits and vegetables. This author proposed the stiffness coefficient of equation (2) to estimate the firmness level of the fruit.

$$firmness_a = f_1^2\, m \tag{2}$$

Where f_1 is the first resonance frequency of the fruit and its mass m. Later, Cooke [7] proposed to estimate the firmness with the stiffness coefficient for spheroidal objects of equation (3).

$$firmness_c = f_1^2\, m^{2/3} \tag{3}$$

These expressions have been extensively used by other researchers to estimate the maturity of fruit using vibrational methods. This parameter, with its two formulations (*2*) and (*3*), have been studied together with the rest of the vibrational parameters to try to predict the presence of internal voids.

The density was estimated from the mass and volume of the watermelon by approximating its shape to that of an ellipsoid according to equation (*4*).

$$volume = 4\frac{\pi}{3}a^2 b \tag{4}$$

Where a and b are the semi-axes of the ellipsoid which are calculated from the equatorial perimeter p_1 and the measured perimeter passing through the pedicle p_2 according to equations (5) and (6).

$$a = \frac{p_1}{2\pi} \tag{5}$$

$$b = \sqrt{\left(\frac{p_2}{2\pi}\right)^2 \cdot 2 - a^2} \tag{6}$$

2.3. Statistical analysis

Firstly, a correlation analysis of the values of mass, density, first peak frequency, estimated firmness according to Abbott and firmness according to Cook was performed. The objective was to find correlations and significant differences between the different parameters, to enable redundant parameters to be discarded and to select the best predictors for the classification algorithms.

After the correlation analysis, the statistical differences between the different physical and vibrational parameters were analysed according to the group of watermelons: without internal defects or with voids. The variance analysis was performed by means of an ANOVA for cases in which the conditions of normality or homoscedasticity were met. For cases where the normality condition was not met, a Kruskal-Wallis test was performed. Normality was tested by means of an Anderson-Darling test and homoscedasticity by means of Bartlett's test. For cases in which correlation was found between different factors, a covariance analysis, ANCOVA, was also performed to thus consider the combined effect of both variables.

After the statistical analysis, the variables with the greatest differences between the two groups of watermelons were selected to be used as predictors in the classification algorithms. To assess the performance of the classifier algorithms, two types of validation were performed: a cross-validation using the same training data and a validation by introducing data from 19 new watermelons into the trained algorithms. The accuracy of different types of classifier algorithms were compared. For this study, the *Classification Learner* software from Matlab [34] was

Table 2
Results of the variance analysis.

	Normality test (p-value)	Homoscedasticity test (p-value)	Variance test type	Analysis of Variance test (p-value)
Mass	0.001	0.911	Kruskal-Wallis	0.043
Radius (*a*)	0.005	0.593	Kruskal-Wallis	<0.001
Volume	0.001	0.548	Kruskal-Wallis	0.002
Apparent density	0.438	0.417	ANOVA	<0.001
f_1	0.064	0.650	ANOVA	<0.001
f_{max}	0.497	0.028	ANOVA	<0.001
$firmness_a$	0.590	0.071	ANOVA	<0.001
$firmness_c$	0.863	0.229	ANOVA	<0.001

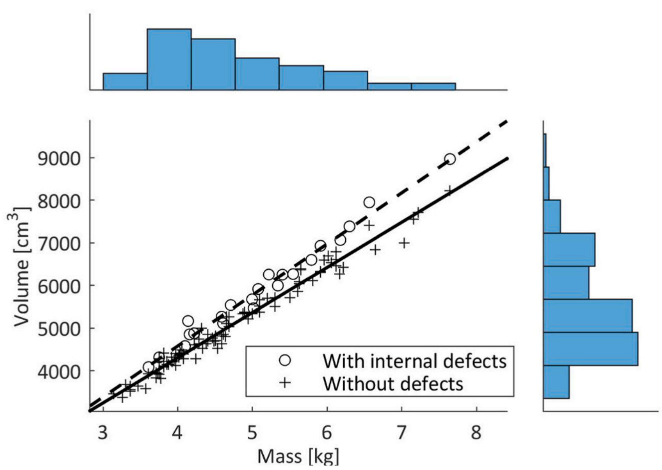

Fig. 4. Masses and volumes for the watermelons tested.

used.

3. Results

3.1. Statistical analysis results

The linear dependence between two parameters was measured using Pearson's correlation coefficient as shown in Table 1. Not surprisingly, mass shows a strong positive correlation with radius. There is also a strong correlation between the two estimators of firmness since they depend on the same properties. The frequency of the first peak and the frequency at which the maximum peak occurs have a weak correlation (Pearson coeff. 0.69) since these two peaks coincide in some watermelons. The weak correlation between f_1 and the calculated Cook firmness estimator is explained by the fact that this frequency is the one used to calculate firmness. Weak correlation between mass and firmness estimators were also found. No correlation was found between mass and density, nor between watermelon mass and the f_1 and f_{max} frequencies.

The results of the analysis of variance are shown in Table 2. The mass and size of the watermelons studied do not display a normal distribution (see Fig. 4) due to the fact that the smallest watermelons are not harvested although the largest ones are.

Both groups of watermelons show statistically significant differences for the parameters considered, with mass showing the smallest differences. On average, watermelons with internal voids show a lower density than watermelons without defects, 0.87 g cm^{-3} and 0.93 g cm^{-3} respectively. However, hollow watermelons are, on average, larger, 5.04 kg and 11.14 cm radius, than watermelon without defects, 4.48 kg and 10.42 cm radius. This reinforces the idea that the presence of internal voids in watermelons is caused by excessively rapid growth.

Taking the covariance into account and using a linear regression model to correct possible differences as a function of the mass variable, by means of an ANCOVA, it was found that, with a 95% significance level, that the mass-adjusted volume values are different for both groups ($p < 0.001$). Fig. 4 shows the regression lines for each group.

Fig. 5 shows the peak frequency values of the transfer function and of the firmness estimators with respect to the mass of each watermelon.

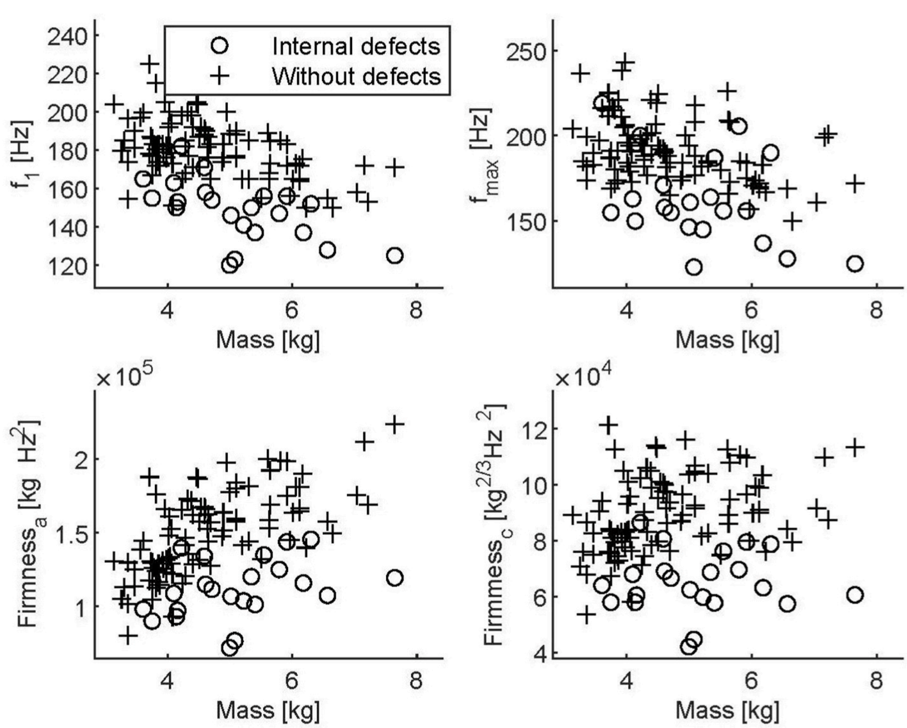

Fig. 5. Mass, firmness estimation and characteristics frequencies f_1 and f_{max} for the watermelons tested.

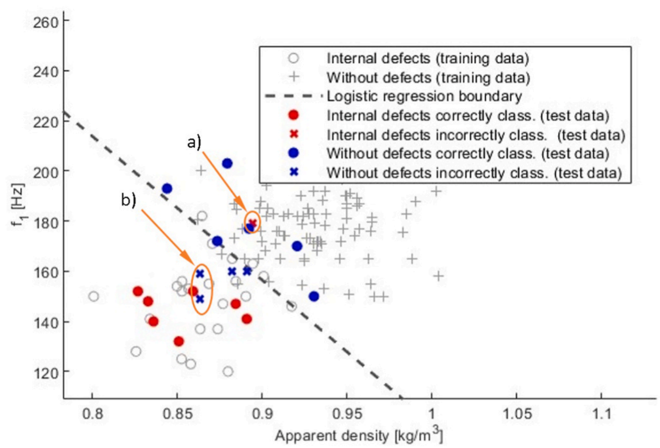

Fig. 6. Classification using logistic regression. Grey: training data; blue and red: validation test data.

Watermelons with internal defects have a lower f_1 value than watermelons without void defects. f_{max} values are more scattered, and the cluster of data points overlap more than in the case of f_1. In view of the results, the frequency at which the first peak occurs in the transfer function seems to be a better predictor of the existence of defects than the frequency at which the maximum occurs in the transfer function. For the firmness values, when using mass and f_1 in both cases, no preference was seen in the use of $firmness_a$ or $firmness_c$.

3.2. Classification algorithms

In the first instance, a logistic regression was tested as a classification algorithm due to its simplicity. Density and frequency f_1 were used as predictors since, in view of previous results, they are those that show the most significant differences. Used with the aforementioned predictors, this algorithm had a 92.2% of accuracy using training data in a five-fold cross-validation. The accuracy of the trained Logistic Regression algorithm classifying the new 19 watermelons test data was 73.68%. Fig. 6 shows the classification boundary of the logistic regression, the training data and validation test data. Only one watermelon with internal defects was classified incorrectly as a "without defects". When this watermelon was opened, the size of the internal defect was found to be very low (see Fig. 7a). There were two other watermelons classified as having internal defects, but upon inspection no holes were found, however, they were found to be overripe. Even with a level of ripeness that makes them unfit for sale, as shown in Fig. 7b (two watermelons with a density of 0.87 kg/m^3). There were another two watermelons without internal voids that were classified as "with internal defects". These watermelons are close to the decision boundary of the logistic regression, see Fig. 6.

Other classification algorithms were also employed using the *Classification Learner* software from Matlab, obtaining the success probabilities shown in Table 3. This table shows the accuracy in a five-fold cross-validation using the training data and the accuracy of the trained algorithms classifying the validation test data.

In view of the results obtained with the logistic regression using two predictors (apparent density and frequency f_1), the validation tests of the classifiers have been carried out using, in addition, Cook's estimation of firmness as a predictor ($firmness_c$). The aim is to evaluate whether the accuracy of the algorithms improves when a third predictor parameter is introduced.

The accuracy of the cross-validation remains between 91 and 94% for all the classifier algorithms shown, remaining the same or slightly higher when introducing the estimated firmness as a third predictor parameter.

The classifier based on neural networks (trilayered) obtained the lowest accuracy (68.42%), followed by logistic regression (73.68%), using only two predictors. These algorithms were also the least accurate when using three predictors (78.95%). Fig. 8a shows the result of the validation of the classifier based on neural networks and using 2 predictors. This algorithm failed in 6 of the 19 watermelons used for the test.

Table 3
Probability of success for most relevant classification algorithm.

Model Type	3 predictors		2 predictors	
	Accuracy (Cross-validation) %	Accuracy (Test) %	Accuracy (Cross-validation) %	Accuracy (Test) %
Discriminant (Quadratic)	94.83	84.21	93.10	78.95
Ensemble (Subspace Discriminant)	93.10	89.47	93.10	84.21
KNN (Cubic)	93.97	89.47	94.83	78.95
Logistic Regression	94.83	78.95	92.24	73.68
Neural Network (Trilayered)	93.10	78.95	91.38	68.42
SVM (Coarse Gaussian)	93.10	89.47	93.10	84.21

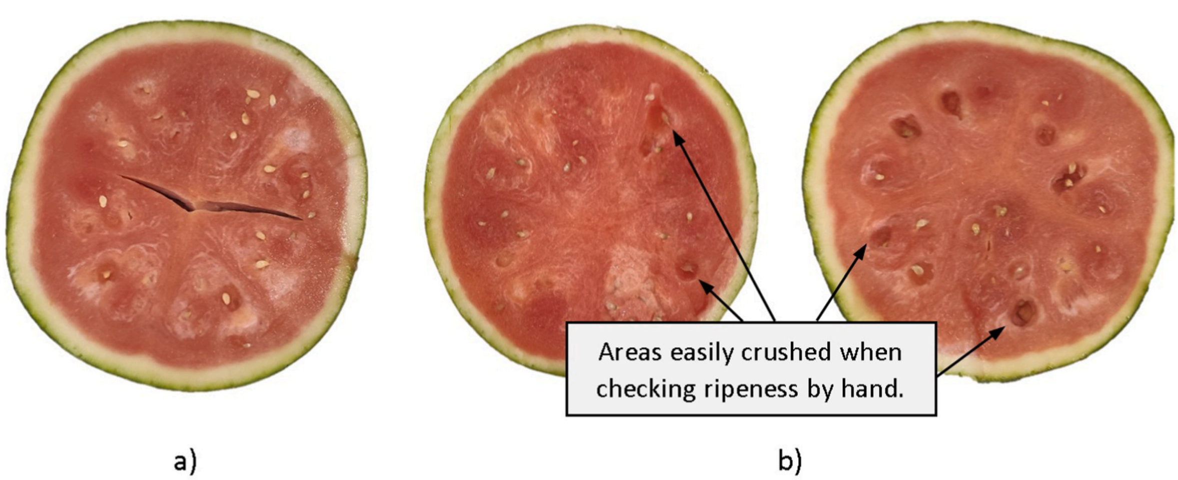

Fig. 7. Watermelons incorrectly classified of Fig. 6 a) watermelon classified as "without defects" with a small hollow; b) watermelons without hollows but very ripe classified as "with internal defects".

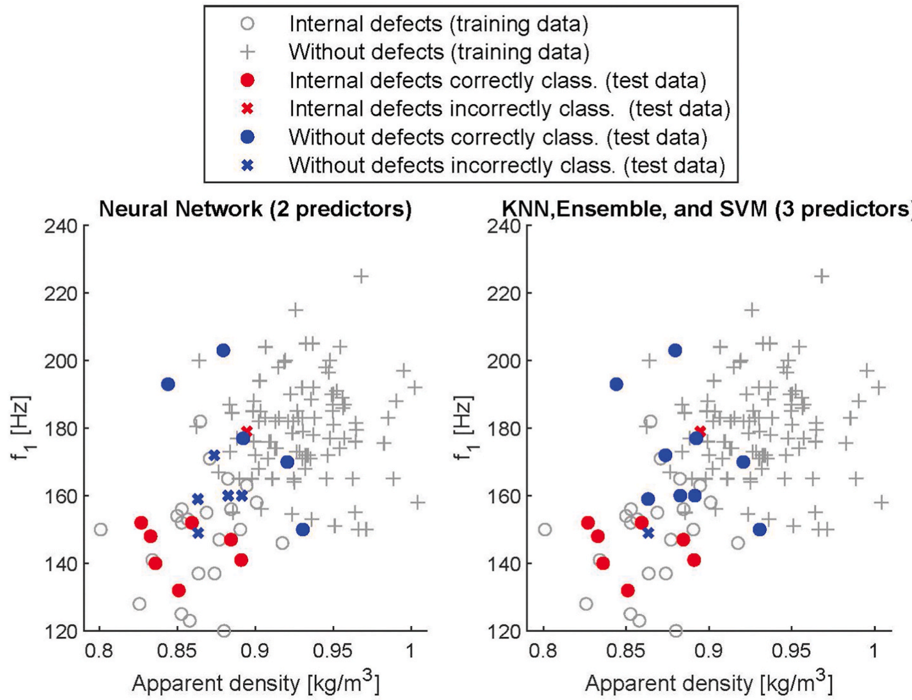

Fig. 8. Classification test results using: a) Neural Network classifier with 2 predictors; b) KNN, Ensemble, and SVM classifier with 3 predictors.

The best accuracy was obtained when using three predictors with classifiers based on KNN, SVM and multiple-classifier techniques (Ensemble) based on discriminant algorithms (89.47%). The use of the single quadratic discriminant classifier achieved an accuracy of 84.21%. The result of the classifiers with the highest accuracy is shown in Fig. 8b. This graph is similar for KNN, Ensemble classifier, and SVM. The algorithms only failed on two of the watermelons mentioned in Fig. 7.

The results shown in Table 3 correspond to the results of the highest accuracy algorithms of each type. A table with the results for other types of classifiers can be found in the appendix. The appendix also includes the results when introducing mass as fourth predictor parameter. But in that case, the classification accuracy did not improve for any of the algorithms with accuracies higher than 78% in the validation test. In fact, the classifier accuracy worsened in many cases when introducing mass as a predictor.

4. Discussion

The analysis of the correlation between watermelon size and peak frequencies showed a similar result to that of studies conducted using vibro-acoustic techniques by other authors: the greater the mass, the lower the resonance frequencies [6,35]. Weak correlation between firmness coefficients and the parameter used to calculate them was found, which agrees with what was found by Taniwaki [36]. When differentiating between watermelons with and without voids, it was found that the frequencies of the first peak and the peak with the greatest amplitude are lower in watermelons with voids, which coincides with other studies [2], and indicates that this vibrational parameter can be used as a predictor of the existence of internal voids.

The presence of voids implies a decrease in watermelon apparent density. No watermelons with internal defects were found with densities higher than 0.92 kg cm^{-3}, which is in line with the studies by Kato [23], which established a density limit of 0.94 kg cm^{-3} for watermelons with voids. However, in our study, watermelons without internal defects were found with densities equal or minor to other watermelons with defects, so it is not possible to classify the watermelons used by exclusively using their density.

Apparent density has proven to be a decisive parameter in the detection of internal voids. For this, the volume of the watermelon must be measured, which, in this study, has been estimated by approximating its shape to that of an ellipsoid of revolution, obtaining good results. There are more accurate methodologies that can be used to determine the volume of the fruit on an industrial scale [37].

The probability of success of the classification algorithms, assessed by cross-validation, was higher than 91% for all types used if the apparent density of the watermelon and the frequency of the first peak of the transfer function are used as predictors. If a watermelon firmness estimator is added as a predictor, the accuracy assessed by cross-validation remains above 93%, improving or remaining the same depending on the type of classifier algorithm used.

The validation test introducing the densities and the first peak frequency of new watermelons shows an accuracy between 68.42%, for a classifier based on Neural Networks, and 84.21% for classifiers based on SVM and Ensemble (Subspace Discriminant). When introducing firmness as a predictor, the accuracy determined by the validation tests increases to 78.95% for the neural network-based classifier and 89.49% for the Ensemble, KNN and SVM classifiers. The success rate is similar to that achieved in other studies using vibro-acoustic techniques (89% achieved by Iglesias et al. [20] and 90% achieved by Noh et al. [21] and Lee et al. [38], and significantly higher than that achieved by experienced humans (71–82%) [39]. Validation has shown that watermelons with thin cracks can be detected as “defect-free”, which has already been noticed by some authors using acoustic methods [38].

Different studies have found that natural frequencies depend on maturity level [7,8,35] and storage time [6]. The maturity level can be estimated by values that depend on the resonance frequencies and mass of the watermelon. Due to their dependence, no reason has been found to use the estimated firmness values instead of frequencies when only 2 predictors parameters are used. When a firmness estimator are introduced into the classification algorithms together with the apparent density and the f_1, the improvement in the success rate is slight. By using a larger number of predictors in the classification algorithms, the results do not improve for those algorithms with success rates above 78%. This suggests that the previous statistical study to discard correlated and redundant parameters is adequate and that the more accurate rating algorithms in Table 3 should be used.

5. Conclusions

Classifiers based on machine learning can be used for the detection of internal voids in watermelons with a high success rate. The apparent density of the watermelon and the first resonance frequency in a slight-impact test are suitable for use as predictors. If a watermelon firmness estimator is also introduced, the probability of the classifier's success rate increases slightly.

The most successful classifiers in validation are those based on SVM, KNN and Ensemble and using three predictors. Success rates of more than 89% have been achieved. Due to the wide variety of size and distribution of internal defects, and ripeness levels, it is difficult to have a totally reliable method and the method presented here should be validated with more samples from future harvests. As several studied have shown the firmness and the extend of internal hollowness are correlated [1], future work should study whether the probability of successful classification is improved by using experimentally measured firmness values, such as by measuring the rigidity of the rind [17,40], or by means of rind colour [5].

In view of the results, it is expected that the methodology applied in this study can be adapted to an industrial environment and can be automated to detect watermelons with voids in a warehouse. It is also expected that this vibration-based technique applied in an automated way will be less costly and more accurate than manual classification and able to compete on cost with other classification techniques [5].

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Francisco Javier Simon Portillo reports equipment, drugs, or supplies was provided by Agrícola Navarro de Haro, S.L.

Data availability

Data will be made available on request.

Acknowledgements

The authors of this paper would like to thank the company, Agrícola Navarro de Haro, S.L., for their collaboration and for providing the watermelons for this study.

Appendix

Table A1Probability of success for each classification algorithm (extended table).

Model Type	4 predictors		3 predictors		2 predictors	
	Accuracy (Cross-validation) %	Accuracy (Test) %	Accuracy (Cross-validation) %	Accuracy (Test) %	Accuracy (Cross-validation) %	Accuracy (Test) %
Discriminant (Linear)	94,83	78,95	93,97	84,21	93,97	78,95
Discriminant (Quadratic)	**93,10**	**84,21**	**94,83**	**84,21**	**93,10**	**78,95**
Ensemble (Boosted Trees)	81,03	52,63	81,03	52,63	81,03	52,63
Ensemble (Bagged Trees)	93,10	78,95	91,38	73,68	91,38	73,68
Ensemble (Subspace Discriminant)	**93,97**	**89,47**	**93,10**	**89,47**	**93,10**	**84,21**
Ensemble (Subspace KNN)	89,66	63,16	87,07	63,16	89,66	73,68
Ensemble (RUSBoosted Trees)	92,24	63,16	91,38	73,68	88,79	68,42
Kernel (SVM)	85,34	63,16	87,93	52,63	81,03	52,63
Kernel (Logistic Regresion)	81,03	52,63	81,90	52,63	81,03	52,63
KNN (Fine)	95,69	78,95	91,38	78,95	88,79	68,42
KNN (Medium)	94,83	89,47	94,83	84,21	93,97	78,95
KNN (Fine)	81,03	52,63	81,03	52,63	81,03	52,63
KNN (Cosine)	93,97	73,68	93,10	73,68	90,52	68,42
KNN (Cubic)	**93,10**	**84,21**	**93,97**	**89,47**	**94,83**	**78,95**
KNN (Weighted)	95,69	84,21	93,97	78,95	90,52	68,42
Logistic Regression	**92,24**	**57,89**	**94,83**	**78,95**	**92,24**	**73,68**
Naive Bayes (Gaussian)	95,69	78,95	96,55	78,95	93,10	78,95
Naive Bayes (Kernel)	94,83	78,95	96,55	78,95	93,97	78,95
Neural Network (Narrow)	93,97	73,68	92,24	73,68	86,21	73,68
Neural Network (Medium)	93,97	73,68	90,52	73,68	90,52	78,95
Neural Network (Wide)	93,97	68,42	91,38	78,95	90,52	73,68
Neural Network (Bilayered)	**93,10**	**73,68**	**91,38**	**73,68**	**88,79**	**78,95**
Neural Network (Trilayered)	**93,10**	**68,42**	**93,10**	**78,95**	**91,38**	**68,42**
SVM (Quadratic)	93,97	73,68	91,38	78,95	90,52	73,68
SVM (Cubic)	93,97	73,68	91,38	68,42	91,38	68,42
SVM (Fine Gaussian)	81,03	63,16	87,07	52,63	85,34	52,63
SVM (Medium Gaussian)	96,55	73,68	94,83	73,68	92,24	73,68
SVM (Coarse Gaussian)	**93,10**	**89,47**	**93,10**	**89,47**	**93,10**	**84,21**
SVM (Linear)	93,10	84,21	93,10	73,68	92,24	78,95
Tree (Fine)	90,52	73,68	89,66	73,68	90,52	73,68
Tree (Medium)	90,52	73,68	89,66	73,68	90,52	73,68
Tree (Coarse)	90,52	73,68	89,66	73,68	90,52	73,68

4 predictors: Apparent density, f_1, fi $rmness_c$, mass.
3 predictors: Apparent density, f_1, fi $rmness_c$.
2 predictors: Apparent density, f_1.

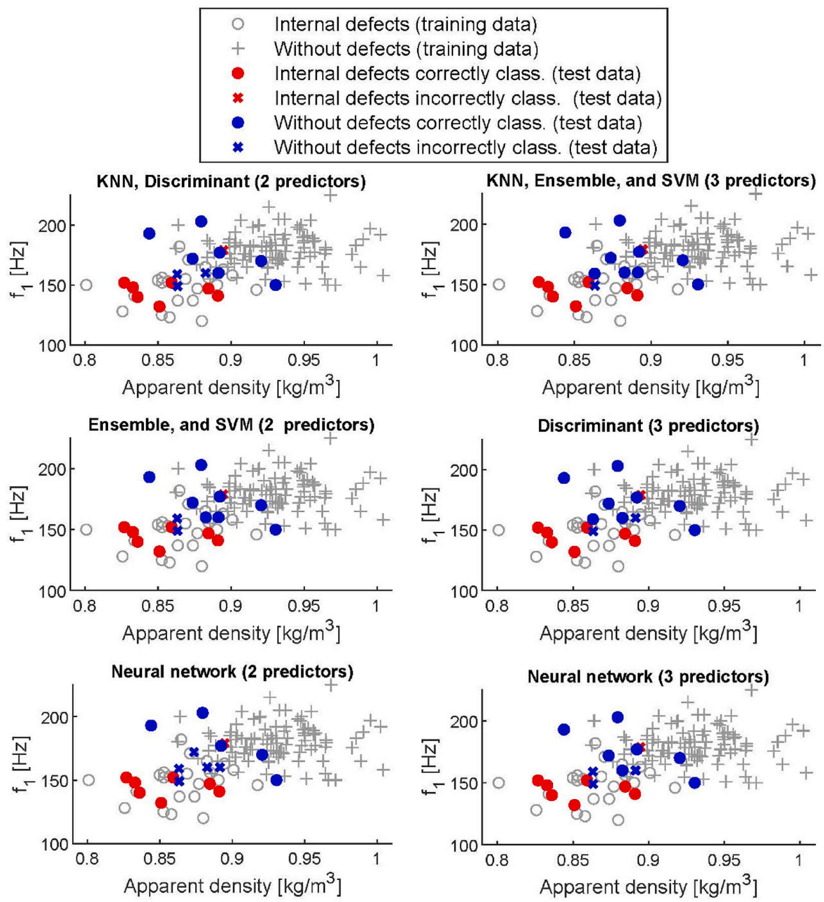

Fig. A1Validation test results for the most relevant classifier algorithms in. Table 3

References

[1] M.A. Trandel, P. Perkins-Veazie, J. Schultheis, Predicting hollow heart incidence in triploid watermelon (Citrullus lanatus), Hortscience 55 (12) (2020) 1926–1930. Dec.

[2] B. Diezma Iglesias, M. Ruiz Altisent, P. Jancsók, Vibrational analysis of seedless watermelons: use in the detection of internal hollows, Spanish J. Agric. Res. 3 (1) (Mar. 2005) 52–60.

[3] M.L. Stone, P.R. Armstrong, X. Zhang, G.H. Brusewitz, D.D. Chen, Watermelon maturity determination in the field using acoustic impulse impedance techniques, Trans. Am. Soc. Agric. Eng. 39 (6) (1996) 2325–2330.

[4] M. Mohd Ali, N. Hashim, S.K. Bejo, R. Shamsudin, Rapid and nondestructive techniques for internal and external quality evaluation of watermelons: a review, Sci. Hortic. (Amst.) 225 (December 2016) (2017) 689–699.

[5] T. Sun, K. Huang, H. Xu, Y. Ying, Research advances in nondestructive determination of internal quality in watermelon/melon: a review, J. Food Eng. 100 (4) (Oct. 2010) 569–577.

[6] H. Yamamoto, M. Iwamoto, S. Haginuma, Acoustic impulse response method for measuring natural frequency of intact fruits and preliminary applications to internal quality evaluation of apples and watermelons, J. Texture Stud. 11 (2) (Jun. 1980) 117–136.

[7] J.R. Cooke, Interpretation of the resonant behavior of intact fruits and vegetables, Trans. Am. Soc. Agric. Eng. 15 (6) (1972) 1075–1080.

[8] J.A. Abbott, G.S. Bachman, R.F. Childers, J. V Fitzgerald, F.J. Matusik, Sonic techniques for measuring texture of fruits and vegetables, Food Technol. 22 (5) (1968) 101–112.

[9] M. Taniwaki, M. Tohro, N. Sakurai, Measurement of ripening speed and determination of the optimum ripeness of melons by a nondestructive acoustic vibration method, Postharvest Biol. Technol. 56 (1) (2010) 101–103.

[10] S. Wang, Q. Jiao, J. Ji, An impulse response method of nondestructive inspection of the ripeness of watermelon, Trans. CSAE 15 (3) (1999) 241–245.

[11] S. Hayashi, J. Sugiyama, K. Otobe, Y. Kikuchi, S. Usui, Nondestructive measurement for maturity of muskmelons by analysis of acoustic signals, Nippon. Shokuhin Kogyo Gakkaishi 39 (6) (1992) 465–470.

[12] N. Ozer, B.A. Engel, J.E. Simon, A multiple impact approach for NON-destructive measurement of fruit firmness and maturity, Trans. ASAE (Am. Soc. Agric. Eng.) 41 (3) (May 1998) 871–876.

[13] J. Nourain, Y. Ying, J. Wang, X. Rao, C. Yu, Firmness evaluation of melon using its vibration characteristic and finite element analysis, J. Zhejiang Univ. - Sci. B 6 (6) (Jun. 2005) 483.

[14] R. Abbaszadeh, A. Rajabipour, H. Sadrnia, M.J. Mahjoob, M. Delshad, H. Ahmadi, Application of modal analysis to the watermelon through finite element modeling for use in ripeness assessment, J. Food Eng. 127 (Apr. 2014) 80–84.

[15] P.R. Armstrong, M.L. Stone, G.H. Brusewitz, Nondestructive acoustic and compression measurements of watermelon for internal damage detection, Appl. Eng. Agric. 13 (5) (1997) 641–645.

[16] D. He, Z. Li, H. Wang, On the characteristics of sound wave forms of watermelons, J. Northwest SCI-TECH Univ. Agric. For. (Nat. Sci. Ed.) (1994) 105–107.

[17] M.I. Al-Haq, J. Sugiyama, A. Tomizawa, Y. Sagara, Nondestructive acoustic firmness tester detects the effect of manure on muskmelon texture, Hortscience 39 (1) (Feb. 2004) 142–145.

[18] A. Mizrach, N. Galili, D.C. Teitel, G. Rosenhouse, Ultrasonic evaluation of some ripening parameters of autumn and winter-grown 'Galia' melons, Sci. Hortic. (Amst.) 56 (4) (Feb. 1994) 291–297.

[19] M.L. Farabee, M.L. Stone, Determination of watermelon maturity with sonic impulse testing, in: International Summer Meeting Sponsored by the, American Society of Agricultural Engineers, 1991. ASAE Paper 91-3013.

[20] B. Iglesias Diezma, M. Ruiz-Altisent, B. Orihuel, Acoustic impulse response for detecting hollow heart in seedless watermelon, Acta Hortic. 599 (Feb. 2003) 249–256.

[21] S.-H. Noh, K.-H. Choi, Non destructive quality evaluation technology for fruits and vegetables, in: International Seminar on Enhancing Export Competitiveness of Asian Fruits, Bangkok, Thailand, 2006.

[22] D. Jie, X. Wei, Review on the recent progress of non-destructive detection technology for internal quality of watermelon, Comput. Electron. Agric. 151 (15) (2018) 156–164.

[23] K. Kato, Electrical density sorting and estimation of soluble solids content of watermelon, J. Agric. Eng. Res. 67 (2) (Jun. 1997) 161–170.

[24] R.E. Garrett, R.B. Furry, Velocity of sonic pulses in apples, Trans. ASAE (Am. Soc. Agric. Eng.) 15 (4) (1972), 0770–0774.

[25] J. Sugiyama, T. Katsurai, J. Hong, H. Koyama, K. Mikuriya, Melon ripeness monitoring by a portable firmness tester, Trans. Am. Soc. Agric. Eng. 41 (1) (1998) 121–127.
[26] S. Ali, K.A. Smith, On learning algorithm selection for classification, Appl. Soft Comput. J. 6 (2) (2006) 119–138.
[27] R. Abbaszadeh, A. Moosavian, A. Rajabipour, G. Najafi, An intelligent procedure for watermelon ripeness detection based on vibration signals, J. Food Sci. Technol. 52 (2) (2015) 1075–1081.
[28] K. Chawgien, S. Kiattisin, Machine learning techniques for classifying the sweetness of watermelon using acoustic signal and image processing, Comput. Electron. Agric. 181 (Feb. 2021), 105938.
[29] A. Alipasandi, A. Mahmoudi, B. Sturm, H. Behfar, S. Zohrabi, Application of meta-heuristic feature selection method in low-cost portable device for watermelon classification using signal processing techniques, Comput. Electron. Agric. 205 (Feb. 2023).
[30] W. Liu, X. Xu, C. Liu, L. Zheng, Rapid discrimination of high-quality watermelon seeds by multispectral imaging combined with chemometric methods, J. Appl. Spectrosc. 85 (6) (2019) 1044–1049.
[31] E.R. Arboleda, K.M. Parazo, C.M. Pareja, Watermelon ripeness detector using near infrared spectroscopy, J. Teknol. dan Sist. Komput. 8 (4) (Oct. 2020) 317–322.
[32] S.L. Agricola Navarro de Haro, Company website, viewed 14 March 2023, [Online]. Available: https://agricolanavarrodeharo.es/en/home/, 2023.
[33] D.J. Ewins, Modal Testing: Theory, Practice and Application, second ed., Wiley, 2009.
[34] MathWorks Inc, Train Models to Classify Data Using Supervised Machine Learning - MATLAB, MATLAB R2022a Help Center, 2022 [Online]. Available: https://www.mathworks.com/help/stats/classificationlearner-app.html. (Accessed 3 October 2022).
[35] J. Mao, Y. Yu, X. Rao, J. Wang, Firmness prediction and modeling by optimizing acoustic device for watermelons, J. Food Eng. 168 (Jan. 2016) 1–6.
[36] M. Taniwaki, M. Takahashi, N. Sakurai, Determination of optimum ripeness for edibility of postharvest melons using nondestructive vibration, Food Res. Int. 42 (1) (2009) 137–141.
[37] G.P. Moreda, J. Ortiz-Cañavate, F.J. García-Ramos, M. Ruiz-Altisent, Non-destructive technologies for fruit and vegetable size determination – a review, J. Food Eng. 92 (2) (May 2009) 119–136.
[38] K. Lee, W. Choi, G. Kim, S. Kang, S. Noh, Internal quality estimation of watermelon by multiple acoustic signal sensing, Key Eng. Mater. 321–323 II (2006) 1209–1212.
[39] B. Diezma-Iglesias, M. Ruiz-Altisent, P. Barreiro, Detection of internal quality in seedless watermelon by acoustic impulse response, Biosyst. Eng. 88 (2) (Jun. 2004) 221–230.
[40] G. Hale, J. Lopresti, D. Stefanelli, R. Jones, L. Bonora, Using non-destructive methods to correlate chilling injury in nectarines with fruit maturity, Acta Hortic. 1012 (Nov) (2013) 83–90.