Comparing quality parameters obtained using destructive and optical methods in grading tomatoes[1]

Comparação de parâmetros de qualidade obtidos por métodos destrutivos e óptico para classificação de tomates

Thainara Rebelo da Silva[2], Anderson Gomide Costa[3]*, Juliana Lobo Paes[3], Marcus Vinícius Morais de Oliveira[3], Francisco de Assis de Carvalho Pinto[4]

ABSTRACT - Optical methods for analysing fruit quality have various advantages compared to conventional methods, including not destroying the sample and the possibility of automating the quality control process. The aim of this study was to compare artificial neural networks developed from biological activity indices obtained using the biospeckle laser optical technique, and from physico-chemical variables obtained by conventional destructive techniques, through an evaluation of their precision in classifying ripe tomatoes, using as a reference an earlier classification carried out by visual inspection. A total of 150 tomatoes were used in the experiment, divided into three ripening stages. Multivariate principal component analysis was used to evaluate interaction of the variance within the groups of data obtained using the biospeckle laser technique and destructive laboratory methods. Two artificial neural networks were developed, the first generated using biological activity indices as input vectors, and the second using physico-chemical variables. The precision of the two neural networks was compared using the Kappa index and overall accuracy, and was based on a reference classification. The variation in ripening as a function of the biological activity indices was explained by the first principal component. The neural network generated from the biological activity indices showed the best performance in classifying the tomatoes into the three ripening stages, with a significant Kappa index and an overall accuracy of 67.5%.

Key words: Biospeckle laser. Artificial neural networks. Physico-chemical variables. Fruit ripening.

RESUMO - A aplicação de métodos ópticos para análise da qualidade de frutos apresenta vantagens em comparação a métodos convencionais, tais como, a não destruição da amostra e a possibilidade de automatizar processos de controle de qualidade. Objetivou-se com este trabalho comparar redes neurais artificiais desenvolvidas a partir de índices de atividade biológica obtidos pela técnica óptica do *biospeckle laser* e a partir das variáveis físico-químicas obtidas por técnicas convencionais destrutivas, avaliando o acerto de classificação quanto a maturação de tomates utilizando como referência uma classificação previamente realizada por inspeção visual. Para realização do experimento foram utilizados 150 tomates divididos em três estádios de maturação. A análise multivariada de componentes principais foi realizada visando avaliar a interação da variância dentro dos grupos de dados obtidos pela técnica do *biospeckle laser e* pelos métodos laboratoriais destrutivos. Duas redes neurais artificiais foram desenvolvidas, sendo a primeira rede neural gerada utilizando os índices de atividade biológica como vetores de entrada e a segunda rede neural gerada utilizando as variáveis físico-químicas como vetores de entrada. A comparação da capacidade de acerto entre as duas redes neurais, tendo como base uma classificação de referência, foi realizada a partir dos parâmetros índice Kappa e exatidão global. A variância da maturação em função dos índices de atividade biológica pôde ser explicada pelo primeiro componente principal. A rede neural gerada a partir dos índices de atividade biológica apresentou melhor desempenho para classificar os tomates nos três estádios de maturação apresentando índice Kappa significativo e uma exatidão global de 67,5%.

Palavras-chave: *Biospeckle laser*. Redes neurais artificias. Variáveis físico-químicas. Maturação de frutos.

INTRODUCTION

The tomato has great economic relevance for Brazilian agriculture, with fruit quality classified and regulated by the Ministry of Agriculture, Livestock and Supply. These regulations are used as a basis for the normative documents produced by quality control agencies for agricultural products (COMPANHIA DE ENTREPOSTOS E ARMAZÉNS GERAIS DE SÃO PAULO, 2003). The insertion of automated systems into agricultural processes is increasingly necessary to obtain higher rates of productivity and to reduce food loss during harvest, post-harvest and marketing (TRENDOV; VARAS; ZENG, 2019).

During tomato ripening, changes occur in some physico-chemical variables that result in changes in the colour, flavour, texture and aroma of the fruit (AHMED *et al*., 2013). In general, conventional methods used to measure physico-chemical variables are destructive, costly, dependent on reagents, and do not allow instantaneous measurement, making it difficult to use this information in automated processes of classification (MUSACCHI; SERRA, 2018; PANDEY; RAVI; CHAUHAN, 2020).

The use of optical techniques is an alternative for inspecting the quality of agricultural products, and helps in monitoring the standards established by quality control agencies (MOHD *et al*., 2017; ZHANG *et al*., 2018). This type of instrumentation is characterised by allowing foods to be analysed - instantly when used in automated systems - without contact and without destroying the sample, and can successfully select fruit based on evaluating its ripeness (COSTA *et al*., 2019), detecting damage (TSOUVALTZIS *et al*., 2020) and predicting physico-chemical attributes (ADEBAYO *et al*., 2016; KUROKAWA *et al*., 2020).

The biospeckle laser is an optical technique that has become established as a tool for analysing the quality of food in general (PANDISELVAM *et al*., 2020; ZDUNEK *et al*., 2014). This technique is based on the optical phenomenon of interference that occurs when a laser beam falls on biological material. When observed successively, the generated patterns are associated with the biological activity found in the analysed material. Among the physical phenomena generated from the interaction of laser light and biological material, light scattering is essential in forming the parameters of the biospeckle laser. The amount of scattering is influenced by the interaction of the characteristics of the light, such as wavelength and frequency, and the cellular structures and metabolic activities of the biological material (BRAGA, 2017).

Several indices have been proposed to quantify biological activity based on measurements of the biospeckle patterns of the laser (ARIZAGA; TRIVI; RABAL, 1999; BRAGA *et al*., 2011; CARDOSO; BRAGA, 2014).

Although there are established applications for these indices individually, their joint use in classification algorithms should be considered, with the aim of improving evaluation of the data obtained with the biospeckle laser. To this effect, supervised classifiers developed from neural networks have been successful in classifying data related to the quality of agricultural products when implemented with the aim of automating systems of fruit selection (COSTA *et al*., 2017a; WU *et al*., 2020).

Therefore, the aim of this study was to compare artificial neural networks developed from biological activity indices obtained using the biospeckle laser optical technique, and from physico-chemical variables obtained by conventional destructive techniques, through an evaluation of their precision in classifying ripe tomatoes, using as a reference an earlier classification carried out by visual inspection.

MATERIAL AND METHODS

A total of 150 tomatoes were used in the experiment, obtained from retail food units located in the district of Seropédica, in the state of Rio de Janeiro, Brazil. First, the longitudinal and transverse diameters of each tomato were measured using digital callipers, and fruit whose shape showed a persimmon-like ratio were selected (COMPANHIA DE ENTREPOSTOS E ARMAZÉNS GERAIS DE SÃO PAULO, 2003). The fruit was divided into three classes according to ripening stage: unripe (green), intermediate (reddish orange) and ripe (predominantly red). Each ripening class consisted of 50 fruits. The fruit comprising each class was selected by visual colour inspection, and grouped according to the standards established by the Companhia de Entrepostos e Armazéns Gerais de São Paulo (2003).

Biospeckle laser images were obtained from each fruit to quantify the biological activity. After obtaining the images, the fruit was submitted to destructive laboratory tests (conventional methods) to obtain the following physico-chemical variables: firmness, weight, pH, total titratable acidity (TTA), total soluble solids (TSS) and water content (WC).

Acquiring the images using the biospeckle laser technique and obtaining the biological activity indices

The apparatus used for acquiring the biospeckle laser images consisted of a high-resolution portable microscope connected directly to the USB port of a computer, a Laserline model iZi 50 mW He-Ne laser, with a wavelength of 655 nm, characterising the colour red, in addition to movable supports, a set of lenses and intensity-reducing filters. The microscope was positioned

to capture the laser light reflected off the fruit, thereby generating the biospeckle pattern (Figure 1). During acquisition of the biospeckle laser images there was no interference from any artificial lighting, as the laser was the only source of light falling on the fruit. Styrofoam plates and foam sheets were placed between the base of the experimental apparatus and the counter where the experiment was carried out to avoid interference from external vibrations at the time the images were acquired. Configuration of the equipment comprising the experimental apparatus remained standardised throughout the experiment.

The Speckle Tools software (GODINHO *et al.*, 2012) was used to acquire the images collected by the portable microscope. During each lighting session, 128 successive 8-bit images relative to the biospeckle patterns were collected at intervals of 0.08 s, with the sampling frequency limited to between 0 and 12.50 Hz. The central lines of the 128 successive images were used to construct a two-dimensional image, known as the Time History Speckle Pattern (THSP), where one dimension of the THSP corresponds to the spatial distribution of the pixels of the biospeckle pattern, and the other dimension corresponds to the variation in pixel intensity over time (ARIZAGA; TRIVI; RABAL, 1999).

From the THSP images, the biological activity of each fruit was quantified using four indices (BAI1, BAI2, BAI3 and BAI4). The BAI1 index was based on the moment of inertia algorithm adapted by Cardoso and Braga (2014) and defined by Equation 1. The BAI2 index was based on the adapted absolute value difference algorithm, and defined by Equation 2. BAI3 was based on the classic second-order statistical algorithm known as the moment of inertia (ARIZAGA; TRIVI; RABAL, 1999), defined by Equation 3. Finally, BAI4 was based on the absolute value difference algorithm proposed by Braga *et al.* (2011) and defined by Equation 4.

$$BAI1 = \frac{0CM(i.j)}{\sum\sum(0CM)}|i-j|^2 \qquad (1)$$

$$BAI2 = \frac{0CM(i.j)}{\sum\sum(0CM)}|i-j| \qquad (2)$$

$$BAI3 = \frac{0CM(i.j)}{Nor.|\sum\sum(0CM)|}|i-j|^2 \qquad (3)$$

$$BAI4 = \frac{OCM(i.j)}{Nor.|\sum\sum(OCM)|}|i-j| \qquad (4)$$

where OCM (i,j) - value of the pixel in row i and column j of the co-occurrence matrix generated from the THSP image, i and j ranging from 0 to 255; Nor. OCM - normalised co-occurrence matrix generated from the THSP image.

Figure 1 - Apparatus for acquiring images from the biospeckle laser and representation of the different ripening stages: (a) unripe, (b) intermediate and (c) mature. (d) Picture of the experimental apparatus
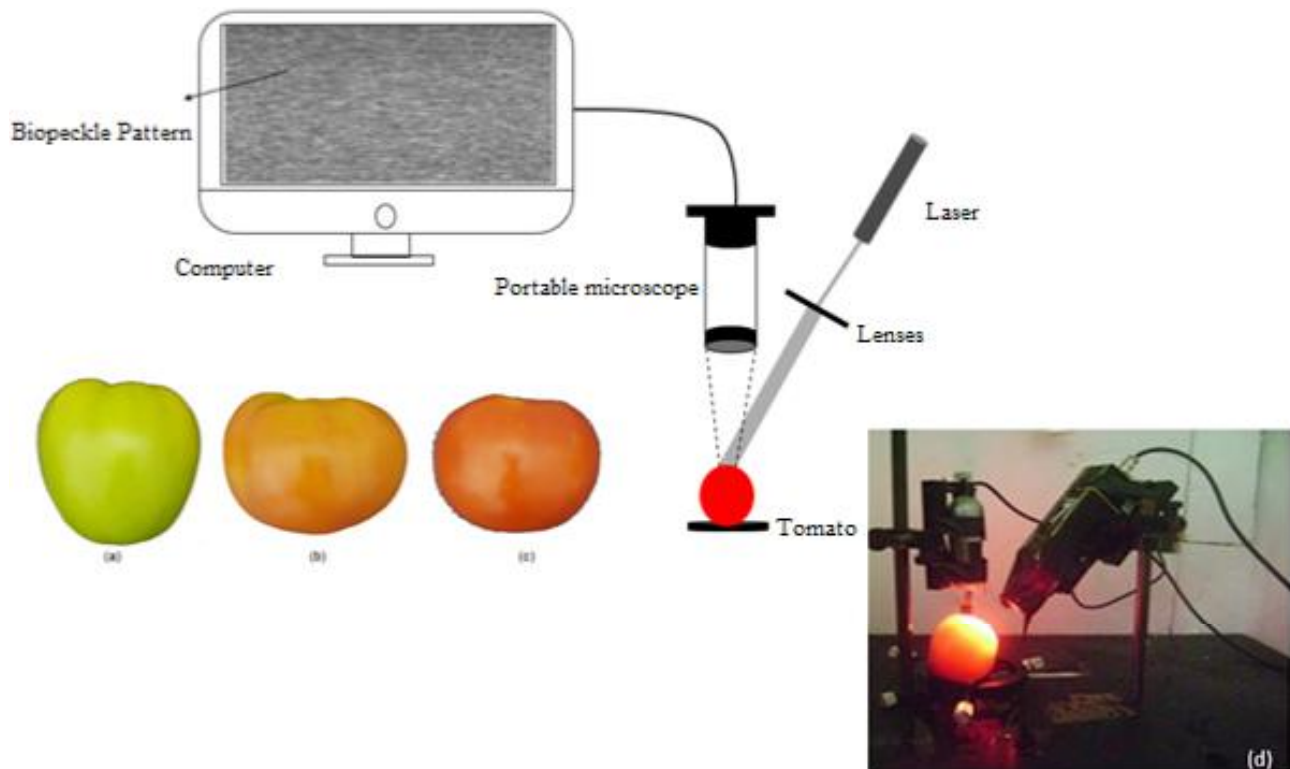
Obtaining the physico-chemical variables

After acquiring the images, the fruit was weighed using a model Ad200 semi-analytical digital balance. The firmness of the pulp of each fruit was determined using an applanater, as per the methodology described by Calbo and Nery (1995).

The chemical variables TSS, pH, TTA and WC were determined following analytical procedures described by the Association of Official Analytical Chemists (2010). The TSS was determined using an Instrutherm, model RTD-95 portable digital refractometer, calibrated with distilled water. The pH was measured by digital pH meter (MS TECNOPON). The TTA was determined using a phenolphthalein indicator applied to 3g of the fruit and titrated in a sodium hydroxide solution. The water content was determined from the ratio between the weight of the water contained in the fruit and the total weight of the fruit (wet weight). The weight of the water was determined as the difference between the total weight of the fruit (wet weight) and the weight of the fruit after 72 h in an oven (dry weight).

Analysing the results

The biological activity indices obtained by the biospeckle laser, and the physico-chemical variables were submitted to descriptive analysis, and the mean, standard deviation, coefficient of variation, and the maximum and minimum values were calculated. The correlation between the parameters was evaluated using the Pearson correlation matrix.

After the descriptive analysis, the data relative to the physico-chemical variables (firmness, weight, pH, TTA, TSS and water content), and the data relative to the biological activity (BAI1, BAI2, BAI3 and BAI4) were submitted to principal component multivariate analysis. The aim of the analysis was to evaluate the interaction of the variables in each data group, and to compare the explanatory power of the variance between the data group obtained by the conventional destructive methods and the data group obtained using the biospeckle laser optical technique.

For the principal component analysis, the data were standardised using the mean and standard deviation (Equation 5), so that the magnitude of the different units would not influence the variance of the sample set.

$$Z_{ij} = \frac{X_{ij} - \overline{X}_j}{S(X_j)} \tag{5}$$

where z_{ij} - standardised value of the variable; X_{ij} - original value of the variable; $\overline{X}_j$ - mean value of variable j; $S(X_j)$ - standard deviation of variable j.

A matrix with the standardised data was created for each data group, with the evaluated parameters arranged in columns and the measurements for each fruit arranged in the rows. The standardised data matrix was used to calculate the covariance matrix between parameters. The eigenvalues were obtained from the linear combination of the covariance matrix, corresponding to the variance of each principal component. For each eigenvalue, an eigenvector was determined, where the values of the eigenvectors corresponded to the coefficients of each principal component.

The percentage of variance explained for each principal component (PC) was evaluated in each data group from the biological activity indices and from the physico-chemical attributes. The correlation was measured to assess the importance of each parameter within each principal component.

Classifying the tomato fruit as a function of ripening using artificial neural networks (ANN)

Artificial neural networks (ANN) were used in a supervised classification of the tomato fruit as a function of the three ripening stages under evaluation. The ANN were developed in the Matlab R2015a® software, using functions from the Neural Network Toolbox (NNTOOL), installed on a computer with an Intel Core i5-5200U 2.20 GHz processor and 8GB of RAM. The aim of the results obtained with the ANN was to compare the classification made with data obtained by conventional methods for evaluating fruit quality (destructive methods) and with data obtained using the biospeckle laser optical technique (non-destructive method). The neural networks used to classify the fruit were of the feedforward type, trained by the error back-propagation algorithm using the Levenberg-Marquardt variation to speed up training and improve performance in classifying the patterns.

In developing the neural networks, the classes were initially identified using a binary numerical system to distinguish the fruit used as supervised samples in training. Fruit grouped by visual inspection in the Unripe class was identified to the classifier as 1-0-0, fruit grouped by visual inspection in the Intermediate class was identified as 0-1-0, and fruit grouped by visual inspection in the Ripe class was identified as 0-0-1.

The architecture of the first neural network (ANN_BAI) comprised an input layer, two intermediate layers and an output layer. The input layer was formed by the characteristic vector, which in this case were the four biological activity indices. The first of the two intermediate layers was composed of four neurons, with the second composed of one neuron. The architecture of the intermediate layers was defined from preliminary tests with different architectures, choosing the neural network that showed the best performance. Finally, the output layer consisted of three neurons, using a binary system for the response, i.e. each neuron generated an output with a value of 0 or 1. Joint analysis of the three output neurons

identified the class in which the fruit was grouped by the classifier. The intermediate layers and the output layers used the hyperbolic tangent as an activation function.

The architecture of the second neural network (ANN_PCV) also comprised one input layer, two intermediate layers and an output layer. The input layer was formed by the characteristic vector, which in this case were the six physico-chemical variables. The first of the two intermediate layers was composed of six neurons and the second layer of four neurons. As before, the architecture of the intermediate layers was defined based on preliminary tests. The output layer was similar to that of the ANN_BAI network, and also employed a binary system to identify the class in which the fruit was grouped by the classifier. The intermediate layers and the output layers also used the hyperbolic tangent as an activation function.

To develop the neural networks, 90 tomatoes (30 fruits from each ripening stage) were randomly chosen, where 80% were for training and 20% for validating the neural network. To test the neural network, the remaining 60 tomatoes were used (20 from each ripening stage). Considering that at the start of training the parameters of each network are randomly generated, and that these values influence training, the architecture was trained 10 times. The neural network that presented the highest percentage success in classifying the test samples was selected.

The performance parameters of the neural network, producer accuracy, user accuracy and overall accuracy, were determined using the confusion matrix obtained in classifying the test samples. The significance of the Kappa index was analysed using the Z-test, and verified whether the classification made by the neural networks could be considered better than a random classification. Since the

fruit was initially allocated to each class based on a visual inspection of the colour, as per standards established by the Companhia de Entrepostos e Armazéns Gerais de São Paulo (2003), the precision and other performance parameters of the classifiers were determined using as a reference the classification previously made by visual analysis.

RESULTS AND DISCUSSION

From the descriptive analysis (Table 1) of the biological activity indices (BAI1, BAI2, BAI3 and BAI4) and the physico-chemical attributes (firmness, weight, pH, TTA, TSS and WC) it can be seen that BAI1 has a higher coefficient of variation, meaning that this index was the most influenced by the different ripening stages. Among the physico-chemical attributes, firmness was the most influenced by the change in ripening stage.

The highest CV (%) found in the values obtained by BAI1 can be explained by the OCM being multiplied by the difference in pixel values (i-j) and raised to the second power, resulting in an increase in the variation of the obtained values. The same occurs in the calculation by BAI3; however, the OCM is divided by the normalised matrix, which generates an increase in scale. As such, these indices tend to present better results when the phenomenon under analysis shows smaller variations, by amplifying the difference in biological activity measured with the biospeckle laser when evaluated at different times (CARDOSO; BRAGA, 2014).

The Pearson correlation (Table 2) between each biological activity index and each physico-chemical attribute of the tomato fruit showed that,

Table 1 - Exploratory analysis of the biological activity indices and physico-chemical variables of the set of 150 tomato fruit

Biological Activity – Biospeckle Laser	Mean	Maximum Value	Minimum Value	Standard Deviation	CV(%)
IAB1	219.00	5400.00	3.00	437.00	200.00
IAB2	10.00	17.00	4.00	3.00	30.00
IAB3	92631.00	271294.00	18549.00	53035.00	57.25
IAB4	2596.00	5329.00	794.00	990.00	38.14
Physico-chemical variables – Laboratory methods	Mean	Maximum Value	Minimum Value	Standard Deviation	CV(%)
Firmness (N.cm⁻²)	5.62	14.28	1.19	2.80	49.82
Weight (g)	159.17	239.78	96.71	28.37	17.82
pH	4.53	4.81	4.35	0.12	2.65
TTA (%)	0.49	0.77	0.26	0.10	20.41
TSS (°Brix)	5.87	6.73	4.70	0.46	7.84
WC (%)	91.31	97.28	60.87	7.28	7.97

TTA – Total titratable acidity; TSS – Total soluble solids; WC – Water content; CV(%) – Coefficient of variation

under the conditions of this research, firmness was the physico-chemical attribute that best correlated with the biospeckle laser, showing significant correlation with each biological activity index. WC and TSS also showed significant correlations with the biological activity indices, with the exception of BAI1.

The correlation between the physico-chemical attributes and the indices generated by the biospeckle laser technique in food has been investigated by various researchers, and gives consistent results, such as when identifying the correlation between biological activity and the TSS, water content and firmness in apples (SZYMANSKA-CHARGOT; ADAMIAK; ZDUNEK, 2012). This relationship is more obvious when there is variation in the physical movement in the interior of the cells due to the production of substances linked to ripening (RETHEESH *et al.*, 2016; ZDUNEK; CYBULSKA, 2011), and to damage (GAO; RAO, 2019) and disease (PIECZYWEK *et al.*, 2018) as a function of the treatment under analysis, which in the case of this research refers to the ripening stage.

Firmness of the pulp is one of the quality parameters that best correlate with measurements taken by the biospeckle laser. Costa *et al.* (2017a) demonstrated that during the weeks of fruit ripening in *Acronomia aculeata*, increases in biological activity are correlated with a reduction in fruit firmness. At the senescence stage, and after the physiological maturation of the fruit, when there is a reduction in fruit firmness over time due to oxidation, there is also a continuous reduction in the biological activity captured by the biospeckle laser.

Considering an accumulated percentage greater than 70% as satisfactory to explain the variability of the data (FERREIRA, 2008), it was concluded from the results shown in (Table 3) that the variation in ripening of the tomatoes, based on the data from the BAI group, can safely be reduced to two PC (accumulated percentage of explained variance, %EV = 96.50%), or even to the first PC (accumulated percentage of explained variance, %EVA > 70.00%). With the data from the PCV group, the variation in ripening has to be explained by using at least the first three principal components (accumulated percentage of explained variance = 70.22%).

Analysis of the correlation of the BAI indices with the three most relevant principal components (Table 4) showed that three indices were strongly

Table 2 - Pearson correlation matrix between the biological activity indices and physico-chemical attributes as a function of the ripening stage of the tomato fruit

	BAI1	BAI2	BAI3	BAI4
Firmness (N cm^{-2})	-0.45*	-0.63**	-0.61**	-0.64**
Weight (g)	0.01^{ns}	0.10^{ns}	0.10^{ns}	0.10^{ns}
pH	-0.02^{ns}	-0.12^{ns}	-0.11^{ns}	-0.11^{ns}
TTA (%)	-0.03^{ns}	0.08^{ns}	0.05^{ns}	0.06^{ns}
TSS (°Brix)	-0.02^{ns}	0.22*	0.20*	0.19*
WC (%)	-0.04^{ns}	-0.52**	-0.52**	-0.50**

TTA – Total titratable acidity; TSS – Total soluble solids; WC – Water content. ** Significant at 1%; * Significant at 5%; ^{ns} Not significant

Table 3 - Percentage of explained variance (%EV) and accumulated percentage of explained variance (%EAV) for each principal component (PC) related to the physico-chemical variables

	BAI		PCV	
PC	%EV	%EAV	%EV	%EAV
PC1	76.13%	76.13%	29.82%	29.82%
PC2	20.37%	96.50%	24.19%	54.01%
PC3	3.07%	99.57%	16.21%	70.22%
PC4	0.43%	100.00%	12.78%	83.00%
PC5	-	-	9.36%	92.36%
PC6	-	-	7.63%	100.00%

correlated with PC1, the only exception being BAI1. Since PC1 showed great capacity for explaining the variance (76.13%), an equation generated from the weights of each variable associated with this component can be used as a quantitative indicator of fruit ripening in the tomato. In analysing the correlation of the PCV with the three most relevant principal components, it was found that firmness, pH and TTA have a stronger correlation with PC1, which indicates that these attributes are the most influenced by the variation in ripening stage of the fruit. On the other hand, TSS and WC are strongly correlated with PC2. In this case, as a single PC did not explain the variance, it was not possible to obtain an equation to indicate ripening of the tomato fruit from a single equation.

The use of physico-chemical variables to distinguish the ripening stage depends on methods that are predominantly destructive, making it difficult to obtain instant results, besides requiring reagents and equipment that generate permanent costs for carrying out the analyses. Therefore, there is a growing demand for research on the use of non-destructive methods, such as the biospeckle laser, whether for analysing fruit quality (ZDUNEK *et al.*, 2014), detecting damage to the fruit (WU; ZHU; REN, 2020) or distinguishing the ripening stage (COSTA *et al.*, 2017b).

The use of the biospeckle laser technique together with an automatic classifier, such as neural networks, is an important step in developing portable sensors that can optimise quality analysis in tomato fruit, and is another tool to help in decision-making for rural producers, industry or consumers.

The earlier classification carried out by visual inspection was used as a reference to compare the precision of the classification made by the neural networks as a function of fruit ripening, using the variables obtained by conventional methods and the indices obtained with the biospeckle laser. The best-performing neural networks obtained from the calibration samples (90 fruits used in this step) were applied to the test sample group (60 fruits used in this step).

The performance parameters obtained from analysing the error matrix generated by the group of test fruit (Table 5), showed that the NN_BAI was more precise in classifying the fruit. The accuracy of the NN_BAI producer showed that 100.00% of the fruit classified as unripe were correctly identified, that 48.60% of the fruit classified as intermediate were correctly identified, and 77.00% of the fruit classified as ripe were correctly identified. For user accuracy, it was found that 50.0% of the fruit classified as ripe (10 fruit) were correctly classified. Of the fruit classified as intermediate, 85.0% (17 fruit) were correctly classified. Finally, 60.0% of the fruit classified as intermediate (12 fruit) were correctly classified. The performance parameters of overall accuracy and the Kappa coefficient showed that the NN_BAI was more precise in classifying the tomato fruit as a function of ripening stage than was the RN_PCV.

The Kappa coefficient of the ANN_BAI was considered 'good' (substantial) according to the scale described in Landis and Koch (1977); the classification made by the neural network was considered better than a random classification at a significance level of 0.01 by Z-test, since a greater value for Z was calculated (Zc) than the tabulated Z value of 2.54. For ANN_PCV, the Z-test proved to be non-significant, indicating that the classification made by the ANN_PCV was equivalent to a random classification. The Kappa coefficient of 0.23 was considered 'reasonably weak' according to the scale described in Landis and Koch (1977).

Table 4 - Correlation between the principal components (PC) and each physico-chemical attribute

	PC1	PC2	PC3
BAI1	0.50	0.86	0.02
BAI2	0.95	-0.11	-0.29
BAI3	0.97	-0.17	0.17
BAI4	0.94	-0.17	0.10
Firmness	0.65	0.24	-0.27
Weight	-0.01	0.37	0.90
pH	0.70	-0.50	0.11
TTA	-0.78	0.25	-0.20
TSS	-0.17	-0.82	0.18
WC	0.49	0.51	-0.04

TTA – Total titratable acidity; TSS – Total soluble solids; WC – Water content

Table 5 - Producer accuracy (PA) and user accuracy (UA) obtained with the confusion matrix for the group of fruit used in the test (20 unripe fruit, 20 coloured fruit and 20 ripe fruit)

	Ripe		Intermediate		Unripe		Performance of the NN		
	PA	UA	PA	UA	PA	UA	OA (%)	Kappa	Z
ANN_BAI	77.0%	50.0%	48.6%	85.0%	100.0%	60.0%	67.5	0.48*	5.35
ANN_PCV	55.6%	25.0%	41.9%	65.0%	55.0%	55.0%	58.1	0.23	2.45

*Significant at 1% by Z-test

Table 6 - Producer accuracy (PA) and user accuracy (AU) defined from the confusion matrix, and performance parameters of the neural network that used the biological activity indices as input vector

	Ripe		Unripe		Performance of the NN		
	PA	UA	PA	UA	OA (%)	Kappa	Z
NN_BAI	85.0%	85.0%	85.0%	85.0%	85.0%	0.70*	6.20

*Significant at 1% by Z-test

From the results obtained by the ANN_BAI, it was found that the use of an artificial visual system based on the principle of the biospeckle laser technique can help in the automated selection of tomato fruit at different ripening stages. The use of the biospeckle laser together with a neural network was also demonstrated by Costa *et al*. (2017b), who obtained a Kappa coefficient of 0.65 and an Overall Accuracy of 82.29% when classifying fruit of the Macaw palm suitable for harvesting.

Fruit at more defined ripening stages (green and ripe) tend to facilitate classification when biological activity indices are used, as these are sensitive to the metabolic changes arising from fruit ripening; however, when these values are close, classification becomes less effective.

Analysing the classification made by the ANN_BAI using the unripe and ripe stages only (Table 6), producer accuracy showed that 85.0% of the unripe fruit (17 fruit) and 85.0% of the mature fruit (17 fruit) were correctly identified. When analysing user accuracy, it was concluded that all the fruit classified as ripe, 85.0% (17 fruit) were correctly classified in the Ripe class. The same user accuracy was also obtained for the fruit classified in the Immature class. Furthermore, there is an increase in the performance parameters, such as overall efficiency, which went from 67.5% to 85.0%, reinforcing the hypothesis that the transition stage (intermediate ripening) influences classification of the tomato fruit.

By providing automatic classification without needing to destroy the samples, application of the biospeckle laser shows that it is a potential tool for obtaining results that may replace or minimise the use of conventional laboratory techniques. However, it should be noted that manipulating and selecting the fruit to be evaluated is a step that requires attention if the analysis is to be successful. In addition to ripening, measured biological activity can be influenced by possible internal damage found in the fruit when the laser light reaches the subsurface layers of the cells, as demonstrated by Gao and Rao (2019). A complementary test, evaluating the influence of internal damage when classifying the fruit should therefore be considered to guarantee the robustness of the classification.

CONCLUSIONS

1. The biological activity indices obtained with the biospeckle laser technique characterised the different ripening stages in tomato fruit, showing a significant correlation with the physico-chemical attributes of firmness, total soluble solids and water content;

2. The use of principal component analysis has made it possible to explain the variance in the physico-chemical attributes by means of the first three components (accumulated percentage of explained variance greater than 70%), while the variance in biological activity was explained by the first principal component.

3. The classifier developed from the neural network that used as an input vector the indices of biological activity measured with the biospeckle laser, showed better performance than the neural network generated from the physico-chemical variables, demonstrating the viability of the optical technique for the automated classification of tomato fruit into different ripening stages.

4. Using tomato fruit at more distinct ripening stages improved the performance parameters of the neural network generated using the biological activity indices as an input vector.

REFERENCES

ADEBAYO, S. E. *et al*. Prediction of quality attributes and ripeness classification of bananas using optical properties. **Scientia Horticulturae**, v. 212, p. 171-182, 2016.

AHMED, L. *et al*. Effect of delactosed whey permeate treatment on physico-chemical, sensorial, nutritional and microbial properties of whole tomatoes during postharvest storage. **LWT - Food Science and Technology**, v. 51, n. 1, p. 367-374, 2013.

ARIZAGA, R.; TRIVI, M.; RABAL, H. Speckle time evolution characterization by the co-occurrence matrix analysis. **Optics & Laser Technology**, v. 31, n. 2, p. 163-169, 1999.

ASSOCIATION OF OFFICIAL ANALYTICAL CHEMISTS. **Official methods of analysis**. Washington: AOAC, 2010. 1094 p.

BRAGA, R. A. When noise became information: state-of-the-art in biospeckle laser. **Ciência e Agrotecnologia**, v. 41, n. 4, p. 359-366, 2017.

BRAGA, R. A. *et al*. Evaluation of activity through dynamic laser speckle using the absolute value of the differences. **Optics Communications**, v. 284, n. 2, p. 646-650, 2011.

CALBO, A. G.; NERY, A. A. Medida de firmeza em hortaliças pela técnica de aplanação. **Horticultura Brasileira**, v. 13, n. 1, p. 14-18, 1995.

CARDOSO, R. R.; BRAGA, R. A. Enhancement of the robustness on dynamic speckle laser numerical analysis. **Optics and Lasers in Engineering**, v. 63, p. 19-24, 2014.

COMPANHIA DE ENTREPOSTOS E ARMAZÉNS GERAIS DE SÃO PAULO. **Normas para classificação de tomates**. São Paulo: CEAGESP, 2003. 2 p.

COSTA, A. G. *et al*. Determination of macaw fruit harvest period by biospeckle laser technique. **African Journal of Agricultural Research**, v. 12, n. 9, p. 674-683, 2017a.

COSTA, A. G. *et al*. Relationship between biospeckle laser technique and firmness of *Acrocomia aculeata* fruits. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 21, n. 1, p. 68-73, 2017b.

COSTA, D. S. *et al*. Development of predictive models for quality and maturation stage attributes of wine grapes using vis-nir reflectance spectroscopy. **Postharvest Biology and Technology**, v. 150, p. 166-178, 2019.

FERREIRA, D. **Estatística multivariada**. Lavras: Editora UFLA, 2008. 662 p.

GAO, Y.; RAO, X. Blackspot bruise in potatoes: susceptibility and biospeckle activity response analysis. **Journal of Food Measurement and Characterization**, v. 13, n. 1, p. 444-453, 2019.

GODINHO, R. P. *et al*. Online biospeckle assessment without loss of definition and resolution by motion history image. **Optics and Lasers in Engineering**, v. 50, n. 3, p. 366-372, 2012.

KUROKAWA, G. *et al*. Indirect determination of moisture using biospeckle technique. **Revista Ciência Agronômica**, v. 51, n. 1, p. 1-5, 2020.

LANDIS, J. R.; KOCH, G. G. The Measurement of observer agreement for categorical data. **Biometrics**, v. 33, n. 1, p. 159, 1977.

MOHD, A. M. *et al*. Optical parameters in food and agricultural processing. **International Food Research Journal**, v. 24, n. 3, p. 908-914, 2017.

MUSACCHI, S.; SERRA, S. Apple fruit quality: overview on pre-harvest factors. **Scientia Horticulturae**, v. 234, p. 409-430, 2018.

PANDEY, A. K.; RAVI, N.; CHAUHAN, O. P. Quality attributes of vacuum fried fruits and vegetables: a review. **Journal of Food Measurement and Characterization**, v. 14, n. 3, p. 1543-1556, 2020.

PANDISELVAM, R. *et al*. Biospeckle laser technique: a novel non-destructive approach for food quality and safety detection. **Trends in Food Science & Technology**, v. 97, p. 1-13, 2020.

PIECZYWEK, P. M. *et al*. Early detection of fungal infection of stored apple fruit with optical sensors: comparison of biospeckle, hyperspectral imaging and chlorophyll fluorescence. **Food Control**, v. 85, p. 327-338, 2018.

RETHEESH, R. *et al*. Use of laser biospeckle for the evaluation of fruit ripening. **Journal of Pure Applied and Industrial Physics**, v. 6, n. 5, p. 65-70, 2016.

SZYMANSKA-CHARGOT, M.; ADAMIAK, A.; ZDUNEK, A. Pre-harvest monitoring of apple fruits development with the use of biospeckle method. **Scientia Horticulturae**, v. 145, p. 23-28, 2012.

TRENDOV, M. N.; VARAS, S.; ZENG, M. **Digital technologies in agriculture and rural areas–Status Report**. Rome, Italy: Food and Agriculture Organization of the United Nations, 2019. 140 p.

TSOUVALTZIS, P. *et al*. Early detection of eggplant fruit stored at chilling temperature using different non-destructive optical techniques and supervised classification algorithms. **Postharvest Biology and Technology**, v. 159, p. 111001, 2020.

WU, A.; ZHU, J.; REN, T. Detection of apple defect using laser-induced light backscattering imaging and convolutional neural network. **Computers & Electrical Engineering**, v. 81, p. 106454, 2020.

ZDUNEK A. *et al*. The biospeckle method for the investigation of agricultural crops: a review. **Optics and Lasers in Engineering**, v. 52, n. 10, p. 276-285, 2014.

ZDUNEK, A.; CYBULSKA, J. Relation of biospeckle activity with quality attributes of apples. **Sensors**, v. 11, n. 6, p. 6317-6327, 2011.

ZHANG, B. *et al*. Challenges and solutions of optical-based nondestructive quality inspection for robotic fruit and vegetable grading systems: technical review. **Trends in Food Science & Technology**, v. 81, p. 213-231, 2018.