Non-destructive Estimation of Chlorophyll *a* Content in Red Delicious Apple Cultivar Based on Spectral and Color Data

Sajad SABZI[a], Yousef ABBASPOUR-GILANDEH[a], Farzad AZADSHAHRAKI[b], Rouhollah KARIMZADEH[c], Elham ILBEYGI[c], Juan Ignacio ARRIBAS[d,e]

[a]*Department of Biosystems Engineering, College of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardabil, IRAN*

[b]*Agricultural Engineering Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, IRAN*

[c]*Department of physics, Shahid Beheshti University, G.C., Tehran 19839, IRAN*

[d]*Department of Teoría de la Señal y Comunicaciones e Ingenieria Telematica, University of Valladolid, 47011, Valladolid, SPAIN*

[e]*Castilla-León Neuroscience Institute (INCYL), University of Salamanca, 37007, Salamanca, SPAIN*

ABSTRACT

Non-destructive estimation of the chemical properties of fruit is an important goal of researchers in the food industry, since online operations, such as fruit packaging based on the amount of different chemical properties and determining different stages of handling, are done based on these estimations. In this study, chlorophyll *a* content in Red Delicious apple cultivar is predicted as a chemical property that is altered by apple ripening stage, using non-destructive spectral and color methods combined. Two artificial intelligence methods based on hybrid Multilayer Perceptron Neural Network - Artificial Bee Colony Algorithm (ANN-ABC) and Partial least squares regression (PLSR) were used in order to obtain a non-destructive estimation of chlorophyll *a* content. In application of the PLSR method, various pre-processing algorithms were used. In order to statistically properly validate the hybrid ANN-ABC predictive method, 20 runs were performed. Results showed that the best regression coefficient of the PLSR method in predicting chlorophyll *a* content using spectral data alone was 0.918. At the same time, the average determination coefficient over 20 repetitions in hybrid ANN-ABC in the estimation of chlorophyll *a* content, using spectral data and color features were higher than 0.92 ± 0.040 and 0.89 ± 0.045, respectively, which to our knowledge is a remarkable non-intrusive estimation result.

Keywords: Spectroscopy; Color features; Non-destructive estimation; Artificial neural network; Regression; Metaheuristic algorithms

1. Introduction

The production and consumption of fruit and vegetables are rising due to their flavor and healthy properties (Li et al 2016; Sabzi et al 2018) There are several methods for grading fruit including grading by the shape, volume, weight, and physicochemical properties (Sabzi et al 2015). The method of grading fruit and vegetables based on physicochemical properties provides a wide choice for consumers (Zhang et al 2018). Estimating these characteristics has already been done using destructive methods, but it is almost impossible to package fruit according to the amount of any physicochemical property, due to the destructive nature of these estimation methods. Nondestructive methods, such as light spectroscopy, have the ability to package fruit and vegetables according to the amount of any physicochemical property (Arendse et al 2018). Other possibility of these non-destructive methods is the recognition of different fruit

varieties, including apple (Eisenstecken et al 2019). Mohammadi et al (2015) used an imaging method to estimate the ripening time of persimmon fruit. In fact, they divide the whole ripening process of this fruit into three stages: unripe, ripe and overripe fruit. Their algorithm performs classification based on the color features of the persimmon skin in two color spaces, RGB and L*a*b*. In order to verify the reliability of the proposed algorithm, they used a number of physical, mechanical, and chemical properties. After several studies, it was found that color features such as R, G, and b* had a significant difference in the various stages of ripening. Finally, two classifications based on Linear Discriminant Analysis and Quadratic Discriminant Analysis (QDA) were used to categorize the ripening stages of persimmon fruit. Results showed that by using QDA, the ripening stages of persimmons were estimated with a precision of 90.24%.

Banana is among most consumed fruit along the globe. Among all food products, banana ranks fourth (Seymour et al 1993). Ripening stage is one of key factors in post-harvest operations and in general fruit processing. This factor determines consumer acceptability and eating quality (Thompson & Burden 1995). Optical methods are a kind of non-destructive methods that have attracted several researcher's interest in recent years (Ali et al 2017; Adebayo et al 2016a). For example, Adebayo et al (2016b) proposed a method for the non-destructive estimation of qualitative characteristics and classification of the banana ripening stages using optical properties. Five wavelengths of 532, 660, 785, 830 and 1060 nm were used in order to predict fruit quality features. Results showed a high correlation between optical properties and banana ripening stages at 532, 660, and 785 nm wavelengths, so that, the different ripening stages of banana fruit were predictable with an accuracy of 53.97%.

Apple is also a very important garden fruit with high demand around the world. For instance, Cardenas-Perez et al (2017) evaluated the estimated physicochemical properties of Malus domestica cv. Golden Delicious apples by a non-destructive method using a computer vision system. 114 apple samples were used to develop the system. Various color spaces were investigated to extract color features, and finally, it was determined that the most suitable color space for measuring the color of the fruit was CIELab. CIELab coordinates include L* (Luminosity), a* (green to red), and b* (blue to yellow). The physicochemical properties extracted from each sample were: titratable acidity, total soluble solids, firmness, and ripening index. Results showed that by combining color features and chemical properties extracted at different stages of apple ripening, a non-destructive evaluation of the fruit would be possible (Cardenas-Perez et al 2017).

Visible/near-infrared hyperspectral imaging (HSI) technique is one of the most effective non-destructive estimation methods for physicochemical properties, followed by an estimation of the fruit ripening stage. In recent years, HSI method has been widely used to assess food quality and safety (Jackman et al 2009; Wu & Sun 2013). Hu et al (2017) investigated the effect of 1-methylcyclopropene (1-MCP), which causes delayed aging of fruit and vegetables, as one of the most effective ethylene reaction inhibitors in preserving the quality of fruit and aggregating the rate of glucose, fructose, and sucrose in different ripening stages of kiwi fruit, in a non-destructive method. In order to test the system, 210 fruit were selected in the range of 0.09-0.13 kg with no visible defects. Vis/NIR laboratory HSI system (Imspector V10E, Spectral Imaging Ltd., Oulu, Finland) was set in reflection mode. The wavelength range of the spectrometer was set at 308-1105 nm, with a resolution of 1002×1004 pixels (spatial spectrum). During the preparation of hyperspectral images, each kiwi fruit was placed on site, and moved by a stepper motor at a speed of 1.3×10^{-3} m s^{-1} and an exposure time of 30×10^{-2} s, under the objective of a CCD camera. Results showed that glucose, fructose, and sucrose values can be properly modeled using spectral data alone.

In recent years, non-destructive methods have been of interest to researchers around the world due to their practical applications in the food industry (predicting the number of physicochemical properties, detection of fruit and vegetables ripening stages, etc.). Therefore, in this study, two non-destructive spectral and color methods combined with Hybrid (ANN-ABC) were used in garden to estimate the amount of chlorophyll *a* as an index of ripening during different growing stages of Red Delicious apple cultivar, due to its skin visible-range color variation in different stages of ripening.

2. Material and Methods

2.1. Sampling

Spectral, color, and chemical data were collected from 56 Red Delicious apple fruit samples that harvested in four different stages of ripening (14 samples of each stage). The first, second, third and fourth stages of harvesting were 135,

145, 155 and 165 days after full blooming. Apples were obtained from Kermanshah, Iran (longitude: 7.03 °E; latitude: 4.22 °N) gardens. Then, they were transferred to Shahid Beheshti University to measure the spectral data, and in order to extract color features and chemical properties of chlorophyll *a*, they were transferred to the Agricultural Engineering Research Institute, Iran.

2.2. Spectroscopy system configuration

Radiation of electromagnetic waves into material causes the absorption and reflection of radiation by material. This interaction of radiation and material leads to the acquisition and presentation of information in spectroscopic studies. In fruit and vegetables, the absorption process of these beams is carried out by the C-H, O-H, and N-H chemical bonds (Nicolaï et al 2007). To do the spectrometry analysis in a proper way, you need to configure the spectroscopy system, for obvious reasons. An Intel Corei3-CFI PC, 330Mb at 2.13GHz, 4GB of RAM and Windows 10 was used, which was also equipped with SpectraWiz software (StellarNet Inc. Tampa, Florida) to store the resulting spectrum in the computer. The measurement mode in this research was reflective mode. The EPP200NIR (StellarNet, Tampa, Florida) model spectrometer with an InGaAs (Indium Gallium Arsenide) detector with operation range of 200 nm to 1100 nm and a resolution of 1 to 3 nm, was used in this study, which was connected to a computer via a USB2 cable. Also, the SLI-CAL (StellarNet, Tampa Florida) model light source made of tungsten halogen with a power of 20 watts was used in this study. A two branch optical fiber was used to guide light from the light source to apple and from apple to the spectrometer. Because of intense noise, the first 200 nm and the last 100 nm wavelengths were eliminated. So, the studied spectral range was 400 nm to 1000 nm. Given the noise in the spectroscopy, 16 samples out of 56 samples were eliminated from the study resulting a total of 40 samples that were used for analysis.

2.3. Extraction of color features

In order to perform a colorimetric study, parameters L*, a* and b* of the skin of apples were measured by a CR-400 (Konika Minolta, Japan) colorimeter and then, the color purity indices (C*) and hue angle (h_a) were calculated using Equations (1) and (2) (Clerici et al 2011):

$$C* = \left[\left(a* \right)^2 + \left(b* \right)^2 \right]^{1/2} \tag{1}$$

$$h_a = \tan^{-1} \left({b*}/{a*} \right) \tag{2}$$

2.4. Measuring true Chlorophyll a content

Due to changes in the color of the fruit during the ripening stages as well as the change in the amount of chlorophyll *a* during this period, we believe that the non-destructive estimation of chlorophyll *a* is very useful, since it is possible to estimate the stage of fruit ripening, specifically Red Delicious apple cultivar (Costa et al 2009; Amoriello et al 2018). Destructive measurement of chlorophyll *a* content was done by the method used in Ncama et al (2017).

Based on this method:
1. 1 gr of rind powder was extracted using 8 mL of 80% acetone.
2. This material was transferred to a glass tube that covered with ice.
3. This glass tube stand for 10 min in this condition.
4. Previous materials were homogenized for 1 min.
5. Centrifuge operation for 10 min at 4 °C.
6. The calculation of pigments based on wavelengths required (Equation (3)).

$$Chl_a = 12.25 A_{663.2} - 2.79 A_{646.8} \tag{3}$$

Where; A-absorbance of a sample at a subscript wavelength, e.g. $A_{663.2}$ is a sample absorbance at 663.2 nm.

2.5. Non-destructive estimation of chlorophyll a *content*

For the non-destructive estimation of chlorophyll *a* content, two spectral and color methods were used. For spectral data, Partial Least Squares Regression (PLSR) method using ParLeS Software (ParLeS_v3.1) (Rossel 2008), and hybrid ANN-ABC method were used. In addition, hybrid ANN-ABC method was also used for color data.

2.5.1. ParLeS software

ParLes is a chemometrics software used for multivariate modeling and prediction. In fact, this software is used for chemical training, research and spectroscopy. This software has the ability of transmitting and pre-processing the received spectrum from different samples by various algorithms (Rossel 2008).

2.5.2. Hybrid artificial neural network-artificial bee colony algorithm (ANN-ABC)

Multi-layer perceptron (MLP) neural network is an effective method for nonlinear modeling of various characteristics. This network has adjustable parameters whose performance depends on the optimal setting of these parameters. Parameters include number of neurons and layers, transfer function, back-propagation training function and back-propagation weight/bias function. In this study, an artificial bee colony algorithm was used to set these parameters. The bee algorithm is an optimization algorithm based on the honey bee swarming behavior suggested by Pham et al (2006). This algorithm is based on the behavior of honeybee in search of food sources. The different stages of the bee algorithm are as follows:

1. Generating initial responses and evaluating them.
2. Selecting better sites (responses) and sending worker bees to those sites.
3. Returning the bees to the hive with artificial dance (producing a neighbor response).
4. Comparing all the bees of a site and select the best case.
5. Replacing non-selected bees with random answers.
6. Saving the position of the best answer - returning to Step 2 if the termination conditions are not met (Pham et al 2006).

In this study, the number of neurons in each layer may be selected in the range of 0-25. The number of layers could be minimum 1 and maximum 3. There were 15 different types of transfer functions, such as *tansig*, *logsig*, *hardlim*, *satlins* and *tribas* to be selected. 19 different back-propagation network training functions such as *trainb*, *traingda*, *traincgb*, *traincgb* and *trainlm* were able to be sub-optimally selected by ABC algorithm. Finally, the back-propagation weight/bias learning function was selected from among 15 different functions such as *learnpn*, *learnk*, *learncon* and *learnwh*. The method is summarized as follows: first, the ABC algorithm considers a similar vector with the mentioned number of parameters, namely a vector with a minimum of 4 and a maximum of 8 members. For example, vector x= [13, 17, *logsig*, *tribas*, *trainb*, *learnpn*] implies a two hidden layers network with 13 and 17 neurons each, *logsig* transfer function in first layer, *tribas* transfer function in second layer, *trainb* back-propagation network training algorithm, and *learnpn* back-propagation weight/bias function. Mean square error (MSE) is a parameter that determines the efficiency of the MLP neural network, which is related to each ABC selected sub-optimal vector sent to it. Such way, each vector with the lowest values of MSE is selected as optimal in order to set the neural network parameters.

2.6. Evaluation parameters of chlorophyll a content predictive model

To evaluate the efficiency of chlorophyll a content estimation models by Hybrid ANN-ABC, the coefficient of determination (R^2), Sum Squared Error (SSE), Mean Absolute Error (MAE), MSE, and Root Mean Square Error (RMSE) were used (Sabzi et al 2013; Sabzi & Arribas 2018). In order to evaluate the predicted models by PLSR, regression coefficient (R_p), adjusted regression coefficient ($R_{p.adj}$), RMSE, Standard Deviation of the Error Distribution (SDE), and Relative Percent Deviation (RPD) were used, (Rossel 2008).

3. Results and Discussion

Figure 1 shows flowchart of the type of data and the used methods to estimation chlorophyll a.
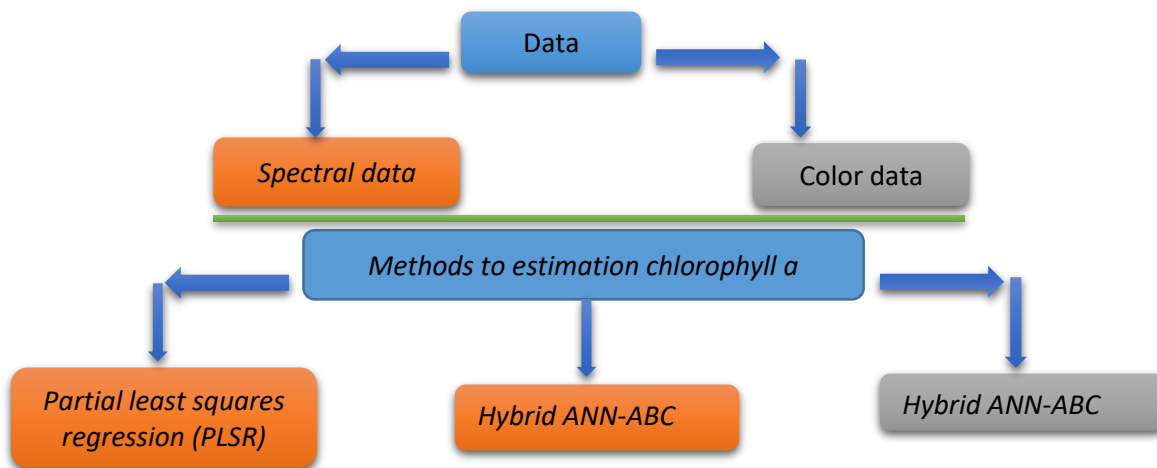
Figure 1- Flowchart of the type of data and the used methods to estimation chlorophyll *a*

3.1. Non-destructive estimation of chlorophyll a content on spectral data using the statistical method

3.1.1. Reflectance and absorption spectra

Nicolaï et al (2007) conducted a study and showed that the reflectance spectra of different fruit such as apples, oranges, nectarines, and pears, were similar. This means that the reasons behind the instance of peaks in the spectral diagrams of fruit are similar. Figure 2 shows the mean reflection and absorption spectra of the visible/near-infrared light. In order to reduce the nonlinearity that may exist in the spectrum, ParLeS software offers the possibility of converting the reflection spectra (R) into the absorption spectra by log (1/R) (Rossel 2008). As you can see, there are peaks near the wavelengths of 490 and 680 nm. A peak near the 490 nm wavelength is related to the absorption of carotenoids and a peak in 680 nm, to the absorption of chlorophyll *a* (Cayuela 2008; Martínez-Valdivieso et al 2014).
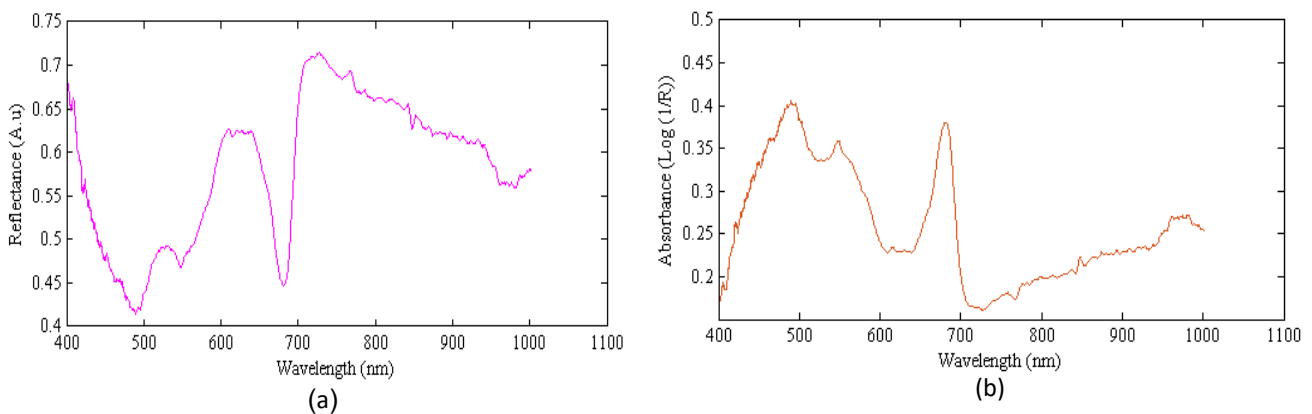

Figure 2- Vis/NIR average spectrum. (a), reflection average spectrum; (b), absorption average spectrum. The unit of reflectance and absorbance is arbitrary unit (A.u)

3.1.2. Pre-processing of spectra

For some reason, such as the effect of light diffusion by changing the detector spacing with the sample, surface roughness in the sample, variation in sample size, noise caused by the increase of the spectrometer temperature, etc., the spectral data, in addition to the sample information, also contains other unwanted information. If one uses ordinary partial least squares regression, this unwanted information reduces the accuracy of the calibration model. On the other hand, ParLeS software, by accessing more stable and reliable models, uses different methods of spectral preprocessing such as Savitzky-Golay (SG), Standard Normal Variate (SNV), Wavelet Filter (WF), SNV with Wavelet Detrending (SNVWWD), Wavelet Detrending (WD), and Median Filter (MF), which are based on various mathematical operations (Rossel 2008).

3.1.3. Partial least squares regression (PLSR) efficiency in predicting the chlorophyll a fruit content

Since the purpose of this study is to estimate the amount of chlorophyll *a* and according to the preceding sections, the light absorption peak near the wavelength of 680 nm is directly related to chlorophyll *a* content, a wavelength window range of 676.282 nm to 686.293 nm was used for estimating Chlorophyll *a* content using PLSR. Table 1 shows the results of the spectral prediction of chlorophyll *a* content using PLSR. In general terms, when Rp, $R_{p.adj}$, and RPD values are large, and at the same time, RMSE Percentage (RMSEP) and SDE values tend to be small, the resulting estimation model will be more efficient and can be used to properly predict the amount of chlorophyll *a*. Based on the explanations given, preprocessing models SNVWW+SG (model 4) and WD+MF (model 5) are better than others and predict chlorophyll *a* content in fruit more accurately, and thus with less error.

Table 1- Calibration results of predictive spectral models of chlorophyll *a* content based on the combination of different preprocessing methods

Model number	Pre-processing	R_p	$R_{p.adj}$	RMSEP (%)	SDE	RPD
1	Without preprocessing	0.727	0.700	0.261	0.255	1.882
2	SNV+SG	0.442	0.386	0.367	0.368	1.335
3	SNV+WF	0.438	0.381	0.368	0.369	1.312
4	SNVWW +SG	0.918	0.910	0.151	0.139	3.204
5	WD + MF	0.918	0.909	0.149	0.141	3.253

3.2. Non-destructive estimation of chlorophyll a content on spectral data using Hybrid ANN-ABC

To predict chlorophyll *a* content by Hybrid ANN-ABC, a spectral window of 662.01 to 698.04 nm, which includes the wavelength of 680 nm, was used. As mentioned above, optimal adjustment of the MLP neural network parameters ensures its high performance in predicting the amount of chlorophyll *a* content. The MLP neural network optimal parameters as set by the artificial bee colony algorithm are two-layer structure with number of neurons 5 and 12. Transfer function for first and second layers are tirbas and purelin respectively. Backpropagation network training function is traingda and final backpropagation weight/bias learning function is learnos. After determining the optimal structure of the artificial neural network, in order to validate the predictive method, 20 replications were performed. The mean±std of R, SSE, MAE, MSE and RMSE were 0.962±0.040, 0.441±0.354, 0.129±0.061, 0.034±0.027 and 0.173±0.064 respectively. Also among these 20 replications, the best values of R, SSE, MAE, MSE and RMSE were 0.987, 0.111, 0.062, 0.008 and 0.092 respectively. As it can be seen, the mean determination coefficient of the 20 repetitions is higher than 0.92, and the standard deviation is 0.04. Therefore, it can be concluded that the performance of Hybrid ANN-ABC is much better than partial least squares regression. Figure 3(a) shows the regression analysis scater plot on the estimated values of chlorophyll *a* content and its true measured values, for the test data. Estimated chlorophyll *a* content is related to the average chlorophyll *a* content of the samples in 20 replicates. As can be seen, the regression coefficient is 0.975. This fact implies that the Hybrid ANN-ABC method shows stability on spectral data.
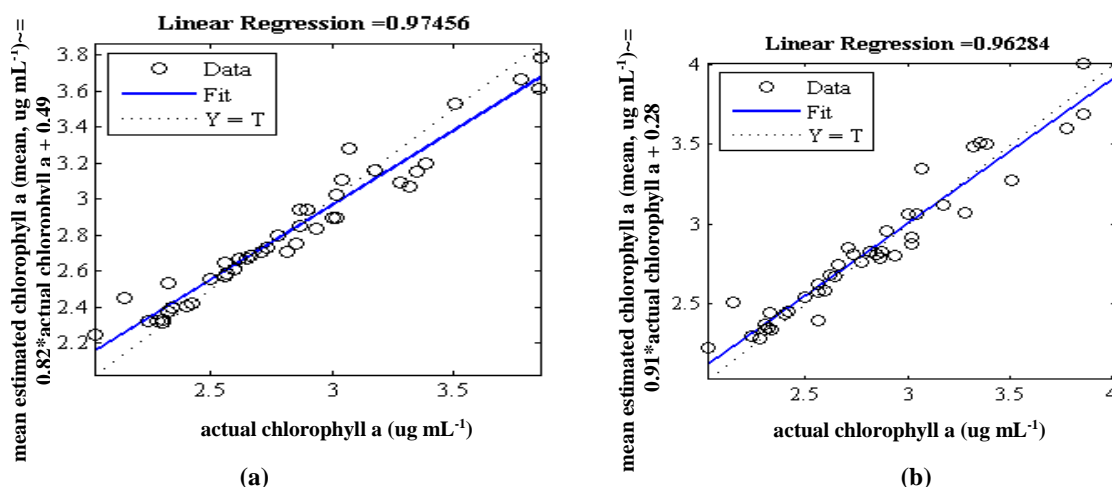

(a) **(b)**

Figure 3- Regression analysis of scatter plot between mean estimated and actual (true measured) apple chlorophyll *a* (ug mL⁻¹) content (test data) using (a), spectral data and (b), color data. Each repetition includes 13 testing samples, so in 20 repetitions 260 samples exist. Since there are 40 input samples in total, there are on average more than 6 different samples over which their average chlorophyll *a* content was measured

3.3. Non-destructive estimation of chlorophyll a content on visible-range color features using Hybrid ANN-ABC

After using different color features as inputs to hybrid ANN-ABC, the results showed that the two properties of the second component of L*a*b* color space (a*) and the hue angle-as input to the artificial neural network-had the higher ability to predict a chlorophyll *a* content as compared to other features. Therefore, these two color properties were used to predict the amount of chlorophyll *a* present in fruit samples. The MLP neural network optimal parameters as set by the artificial bee colony algorithm are two-layer structure with number of neurons 15 and 19. Transfer function for first and second layers are satlin and tirbas respectively. Backpropagation network training function is trainoss and final backpropagation weight/bias learning function is learnwh. In this case, as in the previous section, after determining the optimal amount of artificial neural network parameters, 20 replications were used to determine the stability of the methodology. The mean ± std of R, SSE, MAE, MSE and RMSE were 0.945±0.045, 0.556±0.269, 0.142±0.032, 0.2±0.051 and 0.043±0.021 respectively. Also among these 20 replications, the best values of R, SSE, MAE, MSE and RMSE were 0.982, 0.149, 0.072, 0.107 and 0.011 respectively. As can be seen, the performance of Hybrid ANN-ABC is remarkable in the estimation of chlorophyll *a* content based on color content, but it has a weaker performance than spectral analysis. Figure 3(b) shows the regression analysis of the dispersion plot between the estimated mean and the actual (true measured) value of chlorophyll *a* content in apple (test set) using color data. The regression coefficient of this method over the test data is 0.969, which is an acceptable value in predicting the amount of chlorophyll *a*.

3.4. Comparison of hybrid ANN-ABC method performance in non-destructive estimation of chlorophyll a content using spectral and color data

Figure 4 (a) shows error indices boxplots (MAE, MSE and RMSE) in the Hybrid ANN-ABC method in a non-destructive estimation of chlorophyll *a* content for 20 replicates based on both spectral and color data. As you can see, all the spectral data analysis error boxplot values are lower than the color data counterparts. On the other hand, in general terms the boxplot chart of the mean square error for spectral data is more compact than color data. Latter fact means that the performance of Hybrid ANN-ABC method in the non-destructive estimation of chlorophyll *a* content using spectral data is higher than that based on color data. In a similar fashion as done in Figure 4 (a), Figure 4 (b) also shows the high performance of the Hybrid ANN-ABC method in the non-invasive estimation of chlorophyll *a* content by spectral data. The boxplot charts of regression coefficient and coefficient of determination of the Hybrid ANN-ABC method in the non-destructive estimation of chlorophyll *a* content using spectral data is both tighter and higher than color data counterparts. Partial least squares regression method has poor performance for a state where no pre-processing is performed on the spectral graph. On the other hand, the Hybrid ANN-ABC method on the raw data has a far higher performance than partial least squares regression. This issue is important in real-time applications, since time is a very important factor in real-time applications, and when it is not necessary to preprocess the graph, the estimation of chlorophyll *a* content is possible to be computed in less time. Therefore, the amount of chlorophyll *a* is estimated at a higher speed. The next limiting factor in the non-destructive estimation of chlorophyll *a*, is the cost of building the device. The spectral method is far more expensive than the color method, given that in the color method by only using an algorithm and a typical visible-range camera, it is possible to estimate the non-destructive amount of chlorophyll *a* content. As seen, the performance of the Hybrid ANN-ABC method in the non-destructive estimation of chlorophyll *a* content using color data is close to spectral data. Therefore, considering the cost and performance of the Hybrid ANN-ABC method using color and spectral data, it might be better to use color data in some applications, despite always a trade-off between performance and cost exists. Table 2 shows a comparison of the performance of proposed methods in this study with other non-destructive methods for the estimation of chlorophyll *a* content in the literature. As one can see, the here proposed hybrid ANN-ABC method has better performance than other methods, despite direct comparison is not possible due to the fact that input fruit database is different in each case. It can be concluded that the proposed methodology has proven to have a high-performance and limited cost in the non-destructive estimation of chlorophyll *a* content in Red Delicious apple fruit.

Table 2- Comparison of regression coefficients of different non-destructive estimation methods in estimation of chlorophyll *a* content in various fruit

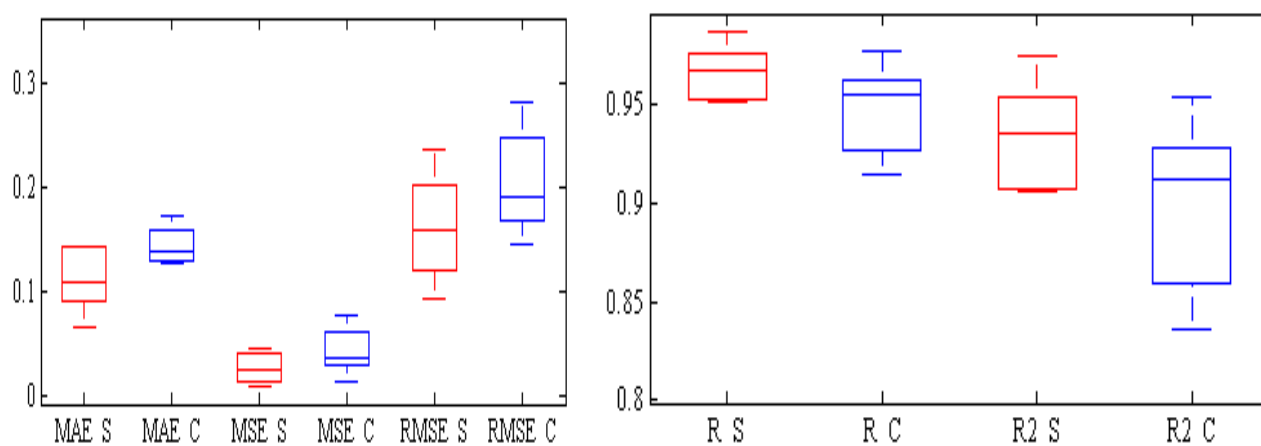| Method	Fruit	Regression coefficient
Proposed method using spectrum data	Apple	0.987
Proposed method using color data	Apple	0.982
(Ncama et al 2017)	Grapefruit	0.943
(Adebayo et al 2016b)	Banana	0.978

Figure 4- Evaluation criteria; (a), boxplot of hybrid ANN-ABC error indices and (b), Regression (R) and determination (R^2) coefficients in non-destructive estimation chlorophyll *a* over 20 runs (test set). S represents spectral data and C represents color data

4. Conclusions

In this study, two partial least squares regression and the Hybrid Artificial Neural Network-Artificial Bee Colony method were used to estimate the amount of chlorophyll *a* as an important parameter correlated with the stage of fruit ripening (Red Delicious apple cultivar in the present study) using both spectral and color data. The most important results are summarized next to conclude:

1. We believe that the performance of the hybrid Multilayer Perceptron Artificial Neural Network - Artificial Bee Colony Algorithm (ANN-ABC) is better than partial least squares regression because of its random nature in the training phase of the predicting model of chlorophyll *a* content.

2. Since the amount of chlorophyll *a* is predictable with rather high-performance by color features and using hybrid ANN-ABC approach, it constitutes an inexpensive method that can be used in on-line conditions that do not need very high accuracy in the estimation.

3. Using spectral data for the non-intrusive estimation of chlorophyll *a* content does not require spectroscopy over the entire visible/near-infrared range, and only a small window around 680 nm wavelength could be used. This will reduce the cost of the configuration and set-up of the spectroscopy system.

4. When using statistical methods like partial least squares regression for non-intrusive estimation of chlorophyll *a* content, the selection the suitable methods for preprocessing of spectral data is important, and ensures their high performance.

Acknowledgements

References

Adebayo S E, Hashim N, Abdan K & Hanafi M (2016a). Application and potential ofbackscattering imaging techniques in agricultural and food processing-a review. *Journal of Food Engineering* 169: 155-164

Adebayo S E, Hashim N, Abdan K, Hanafi M & Mollazade K (2016b). Prediction of quality attributes and ripeness classification of bananasusing optical properties. *Scientia Horticulturae* 212: 171-182

Ali M M, Hashim N, Bejo S K & Shamsudin R (2017). Quality evaluation ofwatermelon using laser-induced backscattering imaging during storage postharvest. *Biology and Technology* 123: 51-59

Amoriello T, Ciccoritti R, Paliotta M & Carbone K (2018). Classification and prediction of early-to-late ripening apricot quality using spectroscopic techniques combined with chemometric tools. *Scientia Horticulturae* 240: 310-317

Arendse E, Fawole O A, Magwaza L S, Nieuwoudt H & Opara U L (2018). Evaluation of biochemical markers associated with the development of husk scald and the use of diffuse reflectance NIR spectroscopy to predict husk scald in pomegranate fruit. *Scientia Horticulturae* 232: 240-249

Cardenas-Perez S, Chanona-Perez J, Mendez-Mendez J V, Calderon-Domı´nguez G, Lopez-Santiago R, Perea-Flores M J & Arzate-Vazquez I (2017). Evaluation of the ripening stages of apple (Golden Delicious) by means of computer vision system. *Biosystem Engineering* 159: 46-58

Cayuela J A (2008). Vis/NIR soluble solids prediction in intact oranges (Citrus sinensis L.) cv. Valencia Late by reflectance. *Postharvest Biology and Technology* 47: 75-80

Clerici M T P S, Kallmann C, Gaspi F O G, Morgano M A, Martinez-Bustos F & Chang Y K (2011). Physical, chemical and technological characteristics of Solanum lycocarpum A. St.-HILL (Solanaceae) fruit flour and starch. *Food Research International* 44: 2143-2150

Costa G, Noferini M, Fiori G & Torrigiani P (2009). Use of vis/NIR spectroscopy to assess fruit ripening stage and improve management in post-harvest chain. *Fresh Produce* 1: 35-41

Eisenstecken D, Stürz B, Robatscher P, Lozano L, Zanella A & Oberhuber M (2019). The potential of near infrared spectroscopy (NIRS) to trace apple origin: Study on different cultivars and orchard elevations. *Postharvest Biology and Technology* 147: 123-131

Hu W, Sun D-W & Blasco J (2017). Rapid monitoring 1-MCP-induced modulation of sugars accumulation in ripening 'Hayward' kiwifruit by Vis/NIR hyperspectral imaging. *Postharvest Biology and Technology* 125: 168-180

Jackman P, Sun D-W & Allen P (2009). Automatic segmentation of beef longissimus dorsi muscle and marbling by an adaptable algorithm. *Meat Science* 83: 187-194

Li J L, Sun D W & Cheng J H (2016). Recent advances in nondestructive analytical techniques for determining the total soluble solids in fruits: A review. *Comprehensive Reviews in Food Science & Food Safety* 15: 897-911

Martínez-Valdivieso D, Font R, Blanco-Díaz M T, Moreno-Rojas J M, Gómez P, Alonso-Moraga Á & Río-Celestino M D (2014). Application of near-infrared reflectance spectroscopy for predicting carotenoid content in summer squash fruit. *Computers and Electronics in Agriculture* 108: 71-79

Mohammadi V, Kheiralipour K & Ghasemi-Varnamkhasti M (2015). Detecting maturity of persimmon fruit based on image processing technique. *Scientia Horticulturae* 184: 123-128

Ncama K, Tesfay S Z, Fawole O A, Opara U L & Magwaza L S (2017). Non-destructive prediction of 'Marsh' grapefruit susceptibility to postharvest rind pitting disorder using reflectanceVis/NIR spectroscopy. *Scientia Horticulturae*

Nicolaï B M, Beullens K, Bobelyn E, Peirs A, Saeys W & Theron I K (2007). Non-destructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology* 46: 99-118

Pham D, Ghanbarzadeh A, Koc E, Otri S, Rahim S & Zaidi M (2006). The bees algo rithm-a novel tool for complex optimisation problems. Paper presented at the *Proceedings of the 2nd Virtual International Conference on Intelligent Production Machines and Systems* (3-14 July 2006), Cardiff, UK

Rossel R A V (2008). ParLeS: Software for chemometric analysis of spectroscopic data. *Chemometrics and Intelligent Laboratory Systems* 90: 72-83

Sabzi S & Arribas J I (2018). A visible-range computer-vision system for automated, non-intrusive assessment of the pH value in Thomson oranges. *Computers in Industry* 99: 69-82

Sabzi S, Javadikia P, Rabani H & Adelkhani A (2013). Mass modeling of Bam orange with ANFIS and SPSS methods for using in machine vision. *Measurement* 46: 3333-3341

Sabzi S, Abbaspour-Gilandeh Y, Javadikia H & Havaskhan H (2015). Automatic grading of emperor apples based on image processing and ANFIS *Tarim Bilimleri Dergisi-Journal of Agricultural Sciences* 21: 326-336

Sabzi S, Abbaspour-Gilandeh Y & Garcia-Mateos G (2018). A new approach for visual identification of orange varieties using neural networks and metaheuristic algorithms. *Information Processing in Agriculture* 5: 162-172

Seymour G B, Taylor J E & Tucker G A (1993). Biochemistry of Fruit Ripening. Springer Netherlands. 107-149

Thompson A K & Burden O J (1995). Harvesting and fruit care. In: Gowen S. (eds) Bananas and Plantains. World Crop Series. Springer, Dordrecht, 84-109

Wu D & Sun D-W (2013). Colour measurements by computer vision for food quality control-a review. *Trends in Food Science & Technology* 29: 5-20

Zhang B, Gu B, Tian G, Zhou J, Huang J & Xiong Y (2018). Challenges and solutions of optical-based nondestructive quality inspection for robotic fruit and vegetable grading systems: A technical review. *Trends in Food Science & Technology*