VISIBLE-NIR INFRARED SPECTROSCOPY FOR POMEGRANATE FRUIT QUALITY ASSESSMENT: CHEMOMETRICS AND COMMON PREPROCESSING METHODS

Rasool Khodabakhshian¹, Bagher Emadi¹*, Mehdi Khojastehpour¹, Mahmood Reza Golzarian¹
¹Department of Biosystem Engineering, Ferdowsi University of Mashhad, P.O. Box 91775-1163, Mashhad, Iran
* E-mail: bagher_emadi@yahoo.com

Abstract
In this study the feasibility of using VIS/NIR spectroscopy along with chemometrics was investigated to predict quality parameters (pH, TSS and firmness) of pomegranate fruit in a nondestructive manner. The effects of different preprocessing methods and spectra treatments, such as column pretreatment (mean centering) and rows pretreatments (including normalization (multiplicative scatter correction (MSC), standard normal variate transformation (SNV)), smoothing (median filter, Savitzky-Golay and wavelet) and transformation (first derivative and second derivative) were analyzed. The results showed that in each studied smoothing techniques SNV gave slightly better results than MSC method. Withal between three studied smoothing techniques namely median filter, Savitzky-Golay and wavelet, the median filter introduced better models. The prediction models were developed by principal component analysis (PCA) and partial least square regression (PLS). The obtained result using first derivative was better for TSS, firmness but second derivative was better for pH. The correlation coefficients (r), RMSEC and RPD for the calibration models were calculated: r=0.95, RMSEC=0.22 °Brix and RPD=6.7 °Brix for TSS; r=0.85, RMSEC=0.068 and RPD=4.58 for pH; r=0.94, RMSEC=0.65 N and RPD=5.65 N for firmness. Also these parameters for the validation models was found to be: r=0.94, RMSEP =0.21 °Brix and RPD=6.72 °Brix for TSS; r=0.86, RMSEP =0.069 and RPD =4.43 for pH; r=0.94, RMSEP =0.68 N and RPD =5.33 N for firmness. It was concluded that VIS/NIR spectroscopy and chemometrics combined with different preprocessing techniques could be an accurate and fast method for nondestructive prediction of key pomegranate quality attributes.

Keywords: VIS/NIR spectroscopy, nondestructive monitoring, pomegranate fruit, quality, preprocessing techniques.


1. INTRODUCTION

The quality evaluation of fruits and vegetables is necessary before any kind of post harvesting processing. The determination of optimum ripening stage is one of the most fundamental aspects that influence on the quality evaluation and depend on a number of internal attributes such as firmness, total soluble solids (TSS) and pH (Moing et al., 1998; Opara, 2000; Nunes et al., 2009; Moghimi et al., 2010). Various methods to measure these attributes are destructive in nature, time consuming and inapplicable to grading and sorting (Salah and Dilshad, 2002; Al-Said et al., 2009; Zarei et al., 2011; Fawole and Opara, 2013a, 2013b). In addition, most of these methods are based on complex processing of sample (Moghimi et al., 2010). So to ensure the minimum acceptability of the quality to consumers, developing efficient and nondestructive methods to measure internal attributes of fruit is important. In recent years, various studied have been reported the development of nondestructive techniques to assess fruit quality. Among them, visible/near infrared (VIS/NIR) spectroscopy seems particularly promising, since it provides fast and reliable information on internal characteristics of many fruit species. The use of NIR to assess the internal
attributes nondestructively has been demonstrated for many agricultural commodities such as apple (Yand et al., 2007; Fan et al., 2009), apricot (Carlini et al., 2000; Camps and Christen, 2009), avocado (Clark et al., 2003), banana (Tarkosova and Copkova, 2000), cherry (Lu, 2001), citrus (Lee et al., 2004; Zude et al., 2008), grape (Herrera et al., 2003; Cao et al., 2010), jujube (Wang et al., 2011), kiwifruit (Moghimi et al., 2010), mandarin (Gomez et al., 2006; Liu et al., 2010), mango (Saranwong et al., 2003), peach (Shao et al., 2011), pear (Xu et al., 2012), pepper (Schulz et al., 2005), plum (Golic and Walsh, 2006), pineapple (Chia et al., 2013), watermelon (Long and Walsh, 2006), and tomato (Shao et al., 2007). In order to respond to the requirement of many parameters such as speed in analysis and flexibility for adapting to different sample states, NIR spectroscopy instrumentation has developed recently. However, many of these instrumentation provides a large amount of spectral data that contains noise and is influenced by a number of physical, chemical, and structural variables (Nicolai et al., 2007; Moghimi et al., 2010; Chia et al., 2013). So, chemometrics (also called Multivariate statistical techniques) are essential to extract the information about quality attributes which is buried in the NIR spectrum. Different integral parts of chemometrics and preprocessing methods such as variable reduction methods, multivariate calibration methods, and different NIR spectroscopy modes including reflectance, interactance and transmission have been investigated by many researchers in order to construct the accurate and reliable models. The main role of spectral preprocessing technique is to remove any irrelevant information which can not be handled properly by the regression techniques. Numerous preprocessing methods have been developed for this purpose such as averaging, centering, smoothing, standardization, normalization and transformation methods (Nicolai et al., 2007). Based on the previous studies, it is clear that the use of incorrect preprocessing methods could cause these methods to perform incorrectly (Nicolai et al., 2007; Moghimi et al., 2010; Chia et al., 2013). Thus, the necessity and the effects of different preprocessing methods should be highlighted. Therefore, the objectives of this study are (i) to investigate the feasibility of using VIS/NIR spectroscopy along with chemometrics to predict quality parameters (pH, TSS and firmness) of pomegranate fruit in a nondestructive manner, and (ii) to study the effects of different combination of preprocessing methods.

2. MATERIALS AND METHODS

Fruit samples
In the present study, a total of 100 pomegranate fruits (ASHRAF variety) without any damage were provided from a commercial orchard in Shahidabad Village, Behshahr County, Mazandaran Province, Iran (Figure 1). All samples were individually washed, labeled and stored in standard refrigeration (3 °C). Before starting the tests, the samples were taken out of the refrigerator and placed under room condition (20 °C and 60% relative humidity) for 2 days to have an equalization room temperature. The samples were randomly divided into two subgroups. The first subgroup of 70 samples was used as a training set for developing partial least square model. The remaining subgroup of 30 fruits was used for model validation and to verify the prediction power of the predictive models. Table 1 shows the summary statistics for some physicochemical properties of samples in each subgroup.

VIS/NIR reflectance spectroscopy collection
From each fruit, four spectra (400-1100 nm at
intervals of 1 nm) in reflectance mode were collected at four equidistance positions along the equator using a dual-channel spectrometer AvaSpec-2048TEC equipped with an AvaSoft7 software for Windows, a cooled, 1 nanometer resolution and sensitivity of 2000 count per 1mJ entrance irradiation in a 0/45° configuration (Fig. 2). The average of these four measurements was used to represent the spectral profile for each sample. The light source consisted of a tungsten halogen lamp (100W, 12V) which is usable in the visible and infrared region. It was arranged at a distance of about 50 mm from the fruit surface and the angle between the incident light source and the fibre optic (that guide reflectance light to a detector) was set to 45 degrees. A white Teflon material was used as the reference material before every measurement. Dark current was measured automatically prior to each measurement. The integration time was set 50ms.

Measuring of quality parameters
After acquiring the spectra, the firmness measurement of samples was made using an Instron Universal Testing Machine (Model H5KS, Tinius Olsen Company) with a 5 mm cylindrical probe programmed to penetrate 8 mm into test fruits with a speed of 10 mm/s. Duplicate puncture tests were performed on opposite sides of equatorial region of each fruit and average value was reported. Peak force required to puncture fruit skin was taken as fruit firmness. Then the samples were macerated with a commercial juice extractor, filtered and centrifuged afterwards. The total soluble solid content (TSS) and pH of juice were measured thrice using a hand-held refractometer (TYM Model, China) and digital pH meter (3020 Model, GenWay Company) respectively, and the average values were noted. All experiments were performed in same conditions.

Chemometrics
Nowadays many researchers found that associated with chemometrics, NIR spectroscopy becomes a powerful tool for the many applications such as agriculture (Roggo et al., 2004), food (Nicolai et al., 2007; Moghimi et al., 2010; Chia et al., 2013), chemical (Larrechi and Callao, 2003), oil industry (Blanco et al., 2001) and pharmaceutical industry (Roggo et al., 2007). Chemometrics is a discipline using mathematical and statistical methods to relate measurements made on a chemical system or process. In other words, the chemometrics regroups several topics such as design of experiments, information extraction methods (Preprocessing, modelling, classification and test of assumptions) and techniques allowing understanding the chemical mechanisms. A review concerning chemometrics has been written by many researchers (Martens and Næs, 1998; Næs et al., 2004; Nicolai et al., 2007), and many textbooks are available (Otto, 1999; Brereton, 2003; Massart et al., 2003). This paper focused on the commonly used preprocessing methods for the analysis of NIR spectra and the calibration models for quantitative and qualitative analysis.

Spectral preprocessing methods
VIS/NIR instruments generate a large amount of spectral data producing valuable analytical information (Blanco and Villarroya, 2002). However, to obtain reliable, accurate and stable calibration models the raw data acquired from spectrometer need to be pre-processed first to reduce the effect of irrelevant information such as background and noisespectra (Cen and He, 2007). The most widely used preprocessing techniques in NIR spectroscopy (in both reflectance and transmittance mode) can be divided into two categories: columns pretreatments and rows pretreatments (Vigni et al., 2013). The column pretreatments include data centering and scaling. The rows
Pretreatments focused on the scattering methods and spectral derivations. These operations are also known as de-noising, smoothing, background and baseline corrections, normalization, alignment (removing horizontal shift), and transformation (Vigni et al., 2013). Firstly in this study, four spectra of every sample were averaged into one spectrum. The averaged value is then converted to absorbance value using Abs=\log (1/R) equation where R is the amount of reflectance, to obtain linear correlation between spectra and sample molecular concentration. Then, several preprocessing methods such as column pretreatment (mean centering) and rows pretreatments (including normalization (multiplicative scatter correction (MSC), standard normal variate transformation (SNV)), smoothing (median filter, Savitzky-Golay and wavelet) and transformation (first derivative and second derivative) were implemented by ParLeS software version 3.1 (Viscarra Rossel, 2008). Centering, which is also referred as mean centering, ensures that all results will be outstanding in terms of variation around the mean (Nicolai et al., 2007). Smoothing is designed to optimize the signal to noise ratio (Nicolai et al., 2007). MSC attempts to remove the effects of scattering by linearizing each spectrum to some ‘ideal’ spectrum of the sample, which, in practice, corresponds to the average spectrum (Nicolai et al., 2007). Also, first and second derivative preprocessing methods were used to remove background spectra and enhance spectral resolution (Cen and He, 2007).

**Calibration and validations**

The preprocessed data were used in the statistical analysis together with the quality parameters. As it was stated earlier, from the 100 spectral samples, 70 were allocated in the calibration set and the remaining 30 were allocated in the validation set. To develop a model between spectral responses of the tested pomegranate fruit and their quality attributes, partial least squares (PLS) regression method was applied to build the model of calibration. The values of one attribute (firmness, TSS and pH) of the calibration set were used to represent the dependent variable (Y). Meanwhile, the reflectance values at studied ranges of wavelengths represented the independent variables or the predictors (X). Because of the vast amount of spectral information provided by NIR spectrophotometers, the large number of samples required to build classification and calibration models, and the high correlation in spectra, there is a need for variable reduction methods that allow the dimensions of the original data to be reduced to a few uncorrelated variables containing only relevant information from the samples. So, before modeling by PLS regression the method of principle component analysis (PCA), a best known and most widely used data reduction method, was employed. The accuracy of the calibration and validation were assessed by correlation coefficient (r), root mean square error of calibration (RMSEC), root mean square error of prediction (RMSEP) and ratio performance deviation (RPD) as follows (Liu et al., 2010b):

\[
 r = \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(\hat{y}_i - y_m)^2}
\]

(1)

\[
 RMSEC = \sqrt{\frac{1}{nc}\sum_{i=1}^{nc}(\hat{y}_i - y_i)^2}
\]

(2)

\[
 RMSEP = \sqrt{\frac{1}{np}\sum_{i=1}^{np}(\hat{y}_i - y_i)^2}
\]

(3)

\[
 RPD = \frac{SD}{RMSEC(P)}
\]

(4)

Where \(\hat{y}_i\) is the predicted value of the i-th observation, \(y_i\) is the measured value of the i-th observation, \(y_m\) is the mean value of the
calibration or prediction set, $n$, $n_c$, and $n_p$ are the total number of observations in the whole data set, the number of observations in calibration and in prediction set, respectively and SD is standard deviation. Generally, a good model should have higher correlation coefficients; lower both RMSEC and RMSEP values, but also a small difference between RMSEC and RMSEP or a RPD value should be more than 5 (Westad et al., 2013).

3. RESULTS AND DISCUSSION

Statistics of the samples

Table 1 shows the summary statistics of some physicochemical properties of samples for both calibration and prediction (validation) data sets. As it can be seen from this table, arils mass and the three main dimensions of arils showed a suitable range. This means that the samples were quite varied in terms of morphology which was the main reason to use normalizing methods (MSC and SNV) and correct themultiplicative and additive effects on the spectra. The range of TSS, pH and firmness were 18.42 to 19.2 °Brix; 3.38 to 3.65 and 38.2 to 42.4 N, respectively.

Vis/NIR spectra of pomegranate fruit

Figure 3 (a) and (b) show the average raw reflectance spectra and absorbance spectrum of Ashraf pomegranate fruit in the wavelength range of 400-1000 nm, respectively. As it was clear in these figures, the spectrum had some absorbance peaks in specific frequencies due to stretching vibration of the overtones of O-H, C-H or N-H functional groups relative to the concentration of some inner compositions with these bands suchas sugars and acids. The absorbance in the range of 400-500 nm was due to the pigments. After 500 nm (in the visible region), the curve had decreasing trend (Figure 3b) and there was a perceptible peak around 750 nm because of the third overtone of O-H and the forth overtone of C-H. Then (in NIR region) the curve had increasing trend and a perceptible peak around 970 nm because of the second overtone of O-H similar to that described by and Gomez et al. (2006) and Cayuela (2008).

Table 1. Statistics of both calibration and prediction data sets for some physicochemical properties of pomegranate fruit samples

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Calibration set (70 samples)</th>
<th>Prediction set (30 samples)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Max</td>
<td>Min</td>
</tr>
<tr>
<td>L* (mm)</td>
<td>66.28</td>
<td>62.03</td>
</tr>
<tr>
<td>W* (mm)</td>
<td>73.15</td>
<td>67.14</td>
</tr>
<tr>
<td>T* (mm)</td>
<td>69.84</td>
<td>67.82</td>
</tr>
<tr>
<td>D_g (mm)</td>
<td>68.25</td>
<td>64.15</td>
</tr>
<tr>
<td>Fruit mass (g)</td>
<td>166</td>
<td>153</td>
</tr>
<tr>
<td>Fruit firmness (N)</td>
<td>41.97</td>
<td>38.5</td>
</tr>
<tr>
<td>Total soluble solid (°Brix)</td>
<td>19.1</td>
<td>18.49</td>
</tr>
<tr>
<td>pH</td>
<td>3.64</td>
<td>3.43</td>
</tr>
</tbody>
</table>

L, length; W, width; T, thickness and $D_g$, geometric mean diameter.
Fig. 1: Fruit and arils of pomegranate (cv. ‘ASHRAF’) cultivar

Fig. 2: a) Setup of Vis/NIR equipment, b) Setup of reflectance mode

Fig. 3: The average raw reflectance Vis/NIR spectra of Ashraf pomegranate fruit (a), and its absorbance spectrum (b)

Table 2. The results of calibration and prediction of PLS models with preprocessing techniques for studied quality parameters
<table>
<thead>
<tr>
<th>Attribute</th>
<th>Preprocessing</th>
<th>No. of PLS factor</th>
<th>Calibration set</th>
<th>Prediction set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>r</td>
<td>RPD</td>
</tr>
<tr>
<td>pH</td>
<td>Original data</td>
<td>15</td>
<td>0.68</td>
<td>2.03</td>
</tr>
<tr>
<td></td>
<td>SNV</td>
<td>12</td>
<td>0.75</td>
<td>3.11</td>
</tr>
<tr>
<td></td>
<td>Median filter</td>
<td>12</td>
<td>0.73</td>
<td>3.05</td>
</tr>
<tr>
<td></td>
<td>SNV Median filter, D^1</td>
<td>9</td>
<td>0.85</td>
<td>4.58</td>
</tr>
<tr>
<td></td>
<td>SNV, SavitzkyGolay, D^1</td>
<td>10</td>
<td>0.81</td>
<td>4.12</td>
</tr>
<tr>
<td></td>
<td>SNV, Wavelet, D^1</td>
<td>10</td>
<td>0.82</td>
<td>4.23</td>
</tr>
<tr>
<td></td>
<td>SNV, Median filter, D^2</td>
<td>8</td>
<td>0.86</td>
<td>5.01</td>
</tr>
<tr>
<td></td>
<td>SNV, Median filter, D^2 and mean center</td>
<td>8</td>
<td>0.84</td>
<td>4.92</td>
</tr>
<tr>
<td></td>
<td>MSC, Median filter, D^1</td>
<td>9</td>
<td>0.84</td>
<td>3.42</td>
</tr>
<tr>
<td></td>
<td>MSC, SavitzkyGolay, D^1</td>
<td>10</td>
<td>0.81</td>
<td>3.21</td>
</tr>
<tr>
<td></td>
<td>MSC, Wavelet, D^1</td>
<td>10</td>
<td>0.79</td>
<td>3.04</td>
</tr>
<tr>
<td></td>
<td>MSC, Median filter, D^1 and mean center</td>
<td>8</td>
<td>0.82</td>
<td>3.35</td>
</tr>
<tr>
<td></td>
<td>MSC, Median filter, D^2 and mean center</td>
<td>9</td>
<td>0.81</td>
<td>2.85</td>
</tr>
<tr>
<td>TSS</td>
<td>Original data</td>
<td>15</td>
<td>0.73</td>
<td>2.21</td>
</tr>
<tr>
<td></td>
<td>SNV</td>
<td>11</td>
<td>0.79</td>
<td>3.36</td>
</tr>
<tr>
<td></td>
<td>Median filter</td>
<td>11</td>
<td>0.78</td>
<td>3.47</td>
</tr>
<tr>
<td></td>
<td>SNV, Median filter, D^1</td>
<td>11</td>
<td>0.95</td>
<td>6.7</td>
</tr>
<tr>
<td></td>
<td>SNV, SavitzkyGolay, D^1</td>
<td>8</td>
<td>0.89</td>
<td>5.06</td>
</tr>
<tr>
<td></td>
<td>SNV, Wavelet, D^1</td>
<td>8</td>
<td>0.89</td>
<td>5.32</td>
</tr>
<tr>
<td></td>
<td>SNV, Median filter, D^1 and mean center</td>
<td>8</td>
<td>0.91</td>
<td>5.01</td>
</tr>
<tr>
<td></td>
<td>SNV, Median filter, D^2 and mean center</td>
<td>9</td>
<td>0.88</td>
<td>4.46</td>
</tr>
<tr>
<td></td>
<td>MSC, Median filter, D^1</td>
<td>8</td>
<td>0.91</td>
<td>4.25</td>
</tr>
<tr>
<td></td>
<td>MSC, SavitzkyGolay, D^1</td>
<td>9</td>
<td>0.88</td>
<td>4.12</td>
</tr>
<tr>
<td></td>
<td>MSC, Wavelet, D^1</td>
<td>7</td>
<td>0.87</td>
<td>4.08</td>
</tr>
<tr>
<td></td>
<td>MSC, Median filter, D^1 and mean center</td>
<td>8</td>
<td>0.9</td>
<td>3.96</td>
</tr>
<tr>
<td></td>
<td>MSC, Median filter, D^2 and mean center</td>
<td>8</td>
<td>0.86</td>
<td>3.53</td>
</tr>
<tr>
<td>firmness</td>
<td>Original data</td>
<td>15</td>
<td>0.69</td>
<td>2.14</td>
</tr>
<tr>
<td></td>
<td>SNV</td>
<td>12</td>
<td>0.78</td>
<td>3.06</td>
</tr>
</tbody>
</table>
Influences of the different studied preprocessing techniques

As it was stated earlier, numerous calibration models were studied by using different preprocessing techniques on the spectral data. In order to investigate the enhanced ability of models based on studied preprocessing techniques, each calibration model was used to predict TSS, pH and firmness of prediction dataset. As it was reported by many researchers a proper model should have a higher correlation coefficients; lower both RMSEC and RMSEP values and also a RPD value more than 5 (Næs et al., 2004; Nicolai et al., 2007; Moghimi et al., 2010; Westad et al., 2013). The results of the most accurate models of calibration and prediction with several preprocessing methods and their combinations for TSS, pH and firmness are summarized in Table 2.

As it can be seen from this Table, the minimum correlation coefficient was found when any preprocessing techniques was not applied for prediction of either TSS, pH or firmness. Though, a decrease in RMSEC and RMSEP values and an increase in correlation coefficient and RPD value was observed when preprocessing techniques was applied. Also it was found that combinations of different preprocessing techniques gave better results to select a proper model. So, in each studied smoothing techniques namely median filter, Savitzky-Golay and wavelet, the PLS model with two different normalizing techniques (MSC and SNV) was evaluated and found that SNV gave slightly better results than MSC method. This is in agreement with the results reported by Moghimi et al. (2010) for SSC and pH prediction of kiwifruit with PLS model preprocessed using MSC and SNV. SNV was applied to remove the multiplicative interferences of scatter, particle size, and the change in light path. MSC was used to compensate for additive (baselineshift) and multiplicative effects in the spectra data which are induced by physical effects. However, the advantage of SNV method over MSC is that SNV is applied to an individual spectrum,
whereas MSC uses a reference spectrum such as the mean spectrum of the calibration set (Nicolai et al., 2007). Withal between three studied smoothing techniques namely median filter, Savitzky-Golay and wavelet, the median filter introduced better models. Also as reported by Moghimi et al. (2010) it is very important to choose the proper median filter rank. In this study the best model was found with a median filter rank of 4. Not only the normalizing method but also the different transformation techniques (first derivative and second derivative) influenced the results for either TSS, pH or firmness.

The obtained result using first derivative was better for TSS, firmness but second derivative was better for pH. Cen and He (2007) have reported that the peaks and troughs were not very noticeable in the original spectra but became more obvious using first derivative. However, these results are in disagreement with the results reported by Liu et al. (2010b) for SSC prediction of navel orange fruit with PLS model and different preprocessing techniques. However, a comparison is not recommended due to the variety of spectrometers and spectral ranges used in these researches. Also as described by Cayuela (2008), the differences between cultivars and varieties such as skin thickness, texture and composition, segment number and seediness may influence spectroscopy measurements.

Fig. 4: Scatter plot of measured versus NIRS predicted TSS for the calibration set (a) and validation set (b) after
SNV, median filter and first derivative

Fig. 5: Scatter plot of measured versus NIRS predicted pH for the calibration set (a) and validation set (b) after SNV, median filter and second derivative

PLS models for quality parameter prediction

Total soluble solids (TSS)

As it is clear from Table 2, the PLS models could predict TSS as well and better than the other studied quality parameters. Many researchers also have reported that NIR spectroscopy with PLS models could predict TSS better than other taste characteristics for various vegetables and fruits such as cherry, mandarin, tomato, and orange fruit (Lu, 2001; Gomez et al., 2006; Shao et al., 2007; Jamshidi et al., 2012). Also, more application of NIR spectroscopy to measure TSS in fruits and vegetables has reported by Nicolai et al. (2007). Withal as it was stated, the PLS model with combination of SNV normalization, median filter smoothing and first derivative for preprocessing can yield better prediction of TSS with $r=0.95$, $RMSEC=0.22$ °Brix and $RPD=6.7$ °Brix. These parameters for the validation models with this combination preprocessing was found to be $r=0.94$, $RMSEP$
=0.21 °Brix and RPD=6.72 °Brix. This is the first reported prediction of TSS in the literatures for pomegranate fruit, to our knowledge, with the best model displaying correlation coefficient and a root mean square error of prediction (RMSEP). Zhang and McCarthy (2013) have reported an \( r=0.41 \) and Root Mean Error of Cross-Validation (RMSECV) 0.57 between measured SSC of pomegranate fruit by the reference analytical methods and predicted SSC by NMR from the PLS model. Figure 4 shows the scatter plot of correlation between the measured and predicted values of TSS for the best model.

**pH**

As it was obvious from Table 2, the normalization, smoothing and transformation methods positively influenced the results of PLS models for pH prediction as SNV was slightly better than MSC when combined with each smoothing and transformation technique. However, the PLS model preprocessed with the combination of SNV, median filter and second derivative with \( r=0.85 \), RMSEC=0.068 °Brix and RPD=4.58 °Brix for calibration set and \( r=0.86 \), RMSEP=0.069 °Brix and RPD=4.43 °Brix for prediction set was preferred.

---

**Fig. 6:** Scatter plot of measured versus NIRS predicted firmness for the calibration set (a) and validation set (b) after SNV, median.
Figure 5 shows the scatter plot of correlation between the measured and predicted values of pH for the best model. As it can be found from this figure and Figure 4, the prediction result of pH was not as accurate as the results of TSS prediction. However, the RMSEC and RMSEP of pH prediction in this research was better. As stated earlier, any published results on the multispectral imaging or NIR spectroscopy of pomegranate fruit and its aril are not available. However, Zhang and McCarthy (2013) found \( r= 0.77 \) and \( \text{RMSECV}= 0.13 \) between measured pH by the reference analytical methods and predicted pH by NMR from the PLS model. They also reported that the RMSECV was very close to RMSEC, which means the loss in the accuracy was very when the calibration models were applied to the test data. In addition they declared that the low value of the error indicated that the PLS model provided fairly accurate prediction of pH.

**Firmness**

For firmness as two other studied quality properties, SNV was preferable against MSC normalizing method to develop the PLS model. Similar to results performance of preprocessing on PLS model for prediction of TSS, the best model for the prediction of firmness was achieved when SNV, median filter and first derivative were used as pre-processing. The scatter plot of the correlation between measured and predicted values of firmness using the best selected model is shown in Fig. 6. Results indicated that VIS/NIR spectroscopy had the potential to predict firmness directly as accurate as it predicts TSS. The obtained statistical parameters for best model were \( r=0.94 \), \( \text{RMSEC}=0.65 \) and \( \text{RPD}=5.65 \) for calibration set and \( r=0.94 \), \( \text{RMSEP} =0.68 \) and \( \text{RPD}=5.33 \) for prediction set. One of important textural property of fruits is firmness that influences the shelf life and consumer acceptance. Firmness shows the chlorophyll and water content of fruits. This is the first reported prediction of firmness in the literatures for pomegranate fruit, to our knowledge.

**4. Conclusions**

This study demonstrate the feasibility of utilizing VIS/NIR spectroscopy and chemometrics combined with different preprocessing techniques to non-destructively characterize key pomegranate quality attributes such as TSS, pH and firmness. Findings indicated that that VIS/NIR spectroscopy with chemometrics is an appropriate tool for nondestructive prediction of internal quality of pomegranate fruit. The prediction models were developed by principal component analysis (PCA) and partial least square regression (PLS). It can be concluded that different pre-processing techniques had effects on the calibration models. The minimum correlation coefficient was found when any preprocessing techniques were not applied for prediction of either TSS, pH or firmness. Though, a decrease in RMSEC and RMSEP values and an increase in correlation coefficient and RPD value was observed when preprocessing techniques was applied. Also it was found that combinations of different preprocessing techniques gave better results to select a proper model. In each studied smoothing techniques SNV gave slightly better results than MSC method. Withal between three studied smoothing techniques namely median filter, Savitzky-Golay and wavelet, the median filter introduced better models. In this study the best model was found with a median filter rank of 4. Not only the normalizing methods and smoothing techniques but also the different transformation techniques (first
derivative and second derivative) influenced the results for either TSS, pH or firmness. The obtained result using first derivative was better for TSS, firmness but second derivative was better for pH. The correlation coefficient (r), RMSEC and RPD for the calibration models was found to be: r=0.95, RMSEC=0.22 °Brix and RPD=6.7 °Brix for TSS; r=0.85, RMSEC=0.068 and RPD=4.58 for pH; r=0.94, RMSEC=0.65 N and RPD=5.65 N for firmness). Also these parameters for the validation models was found to be: r=0.94, RMSEP =0.21 °Brix and RPD=6.72 °Brix for TSS; r=0.68, RMSEP=0.069 and RPD=4.43 for pH; r=0.94, RMSEP =0.68 N and RPD =5.33 N for firmness.

5. ACKNOWLEDGMENTS

The authors would like to thank the Ferdowsi University of Mashhad for providing the laboratory facilities and financial support through the project No. of 28580.

6. REFERENCES


