Quantitative evaluation of impact damage to apples using NIR hyperspectral imaging

Peipei Zhang\textsuperscript{a}, Hongwei Ji\textsuperscript{a,b}, Huaiwen Wang\textsuperscript{b}, Yuexuan Liu\textsuperscript{b}, Xiaochuan Zhang\textsuperscript{b}, and Chunhua Ren\textsuperscript{b}

\textsuperscript{a}School of Mechanical Engineering, Tianjin University, Tianjin, China; \textsuperscript{b}Tianjin Key Laboratory of Refrigeration Technology, Tianjin University of Commerce, Tianjin, China

\begin{abstract}
Impact damage to apples is one of the most crucial quality factors and needs to be detected in postharvest quality sorting processes. In this study, the impact damage of the 'Red Fuji' apple fruit was investigated quantitively by hyperspectral imaging technology. A total of 240 samples were prepared with six groups for different damage degrees. The hyperspectral imaging technique based on near-infrared (NIR) spectrometry in the range of 900–1700 nm was used to measure mechanical parameters, such as the average pressure, contact load, damaged area, absorbed energy, and damaged firmness. Four types of spectral pre-treatment, including the standard normal variate, multiplicative scatter correction, first-order derivative, and second-order derivative, were adopted to improve the model's predictive performance. The quantitative relationships between spectra and mechanical parameters were successfully modeled based on partial least squares (PLS) regression. For 'Red Fuji' apples, raw spectral data without pre-treatment performed better than those after spectral pre-treatments. In this model, the characteristic wavelengths were selected by the Savitzky–Golay second-order derivative (SG 2\textsuperscript{nd} Der) and competitive adaptive reweighted sampling (CARS) method. The results indicate that the CARS-PLS regression model produced better results than the SG 2\textsuperscript{nd} Der-PLS regression model. The good prediction performances were presented by the coefficient of determination ($R_0^2$) and root mean square errors of prediction (RMSEP) values. The $R_0^2$ and RMSEP results of the average pressure, contact load, damaged area, absorbed energy, and damaged firmness are 0.66 and 0.02 MPa, 0.86 and 53.80 N, 0.83 and 116.37 mm\textsuperscript{2}, 0.81 and 0.24 J, and 0.64 and 0.19 N, respectively. This study demonstrates the potential of the NIR hyperspectral imaging technique as a highly accurate way to quantitatively predict the mechanical parameters of apples.
\end{abstract}

\section*{Introduction}

The apple is one of the most popular and highest-selling fruit among consumers due to its rich nutrients, appealing flavor, and crunchy texture. According to a survey of the National Bureau of Statistics, China produced 39.23 million tons of apples, with an export value of 1298.93 million dollars, in 2018. However, apples are inevitably affected by various kinds of external forces during the process of harvesting, packaging, storage, and transport.\textsuperscript{[1,2]} The economic loss caused by the mechanical loss of fruit accounts for about 30\% of the total weights. Mechanical damage is a primary cause of quality loss and degradation for apples according to investigations and the most important cause of bruising is...
conducted by excessive impact forces. Due to the red skin of Fuji apples, it is difficult to detect the bruise damage which took place beneath the peel at an early stage. Economic losses cannot be accurately estimated due to the lack of an approach for an objective quantitative evaluation of the degree of apple damage. Therefore, it is very significant to propose an effective method for quantifying the degree of apple fruit damage caused by external forces.

Numerous methods have been applied to study and detect fruit damage in recent years, such as nuclear magnetic resonance, X-ray computed tomography, and optical coherence tomography images. These technical instruments are promising, but relatively expensive and costly, making it difficult to realize industrial applications at the current technological development. Hyperspectral imaging is a rapid, nondestructive, and non-contact technique. It integrates spectroscopy and imaging techniques, and provides spectral and spatial information simultaneously, which originates from the fields of chemistry and remote sensing. According to recent research results, it has been widely applied to detect the quality of agricultural products, which is an important means of food quality and safety monitoring. For example, it can detect bruise and fungi contamination in strawberries, the hollowness classification of white radish, the common defects on oranges, chilling injury in cucumber fruit. In order to detect the bruising and damage to fruit, many studies have given a variety of fruit classification or prediction models by relating the wavelength to damage features using hyperspectral imaging. At present, most of these studies focus on the distinguishing whether there is damage or not and early detection of fruit damage. Meanwhile, the use of hyperspectral imaging to quantify the impact damage to apples is rarely reported. A quantitative prediction of apple impact damage is achieved based on hyperspectral imaging technology and combined with high-speed camera and pressure-sensitive film technology in this study. At the same time, the potential of hyperspectral imaging technology is confirmed for predicting mechanical parameters of apples.

Therefore, the major objective of this study was to investigate the feasibility of predicting mechanical parameters on damaged ‘Red Fuji’ apples using hyperspectral imaging technique. Four types of preprocessing methods were used to process raw hyperspectral data, and the characteristic wavelengths were selected based on the SG 2nd Der and CARS method. The degree of apple damage was assessed quantitatively by establishing the PLS regression model between the measured spectra and mechanical parameters. It is hoped that the results can provide a theoretical basis for predicting the damage of fruit and helping to reduce economic losses.

Materials and methods

Apple samples and the drop test

‘Red Fuji’ apples were purchased from a local market in Tianjin, China. Samples were stored in conditions with a room temperature of 20°C and 39% relative humidity, in order to decrease the effect of temperature. To ensure the reliability of the work, a total of 240 regular-shaped apples free of visual defects and bruises were selected, and had a height of about 7.5 cm, equator diameter of approximately 8 cm, and mass of 185 g ± 18 g. Samples were divided into six groups on average, with each group containing 40 apples. One group was regarded as the control group to obtain the raw spectral data without impact damage. The other five groups were used as the test group to drop samples from different heights in order to obtain different impact degrees. Forty samples were dropped from the same drop height. The five groups of samples were named H0.3, H0.6, H0.9, H1.2, and H1.5, corresponding to the five drop heights of 0.3, 0.6, 0.9, 1.2, and 1.5 m, respectively.

The drop process of apples was conducted by a drop test machine (PD-315, Suzhou New District Dongling Vibration Testing Instrument Co., Ltd.). Due to the destruction of the peel tissue, the damaged area became soft. The boundary of the damaged area was marked for hyperspectral image acquisition.

The pressure-sensitive film was placed on the steel plate of the drop test machine to measure the contact stress distribution. After preliminary experiments, the ultra-super-low pressure film was selected for this experiment, with a measurement range of 0.2 ~ 0.6 MPa and a test accuracy of ±
10%. The film consisted of two coating sheets, including a sheet with color-forming material and a sheet with color-developing material. The two coating surfaces were placed so that they were facing each other during the test. When pressure was applied to the film, the microcapsules in the pressure area were ruptured to release the color-forming material, which was then absorbed by the color-developing material and a chemical reaction occurred, showing red with different concentrations. A pressure measurement and analysis system (FPD-8010E, Fuji Film Corporation, Japan) were used to evaluate the pressure results. The film was scanned by FPD-8010E to perform numerical analysis, in order to obtain the parameters such as the contact stress peak value, mean value, and stress distribution area.

During the drop test, the falling and rebounding process of apples was recorded by a high-speed motion camera. Apples’ spatial location was recorded at a speed of 2000 f/s with resolutions of 1024 pixel × 1024 pixel. The flow chart employed for evaluating the mechanical parameters is shown in Figure 1.

Absorbed Energy

The digital image correlation (DIC) technique\textsuperscript{[18–20]} is widely accepted and commonly used in many fields for displacement field measurement and strain field estimation. In this study, three or four contiguous images of apples at the moment of drop and the rebound were selected to calculate the displacement. Therefore, the speed of apples can be calculated by:

\[ v = \frac{s}{t} \]  

where \( s \) means the apple’s displacement and \( t \) means the time between selected adjacent images, which was \( 1/2000 \) s. The absorbed energy was calculated by subtracting the rebound energy from the impact energy, as shown in Eq. 2:

---

**Figure 1.** Flow chart of the main steps employed for evaluating the mechanical parameters of apples using hyperspectral imaging.
\[
E_{\text{imp}} = E_{\text{kin}(t-1)} \\
E_{\text{reb}} = E_{\text{kin}(t+1)} \\
\{ \begin{align*}
E_{\text{ab}} &= E_{\text{imp}} - E_{\text{reb}} \\
E_{\text{kin}(t-1)} &= \frac{1}{2}mv_{t-1}^2 \\
E_{\text{kin}(t+1)} &= \frac{1}{2}mv_{t+1}^2
\end{align*} \] (2)

where \(E_{\text{imp}}\), \(E_{\text{reb}}\), and \(E_{\text{ab}}\) represent the impact energy, rebound energy, and absorbed energy for damage, respectively. \(v_{t-1}\) and \(E_{\text{kin}(t-1)}\) represent the velocity and kinetic energy of the last moment before the apple’s impact on the steel plate, and \(v_{t+1}\) and \(E_{\text{kin}(t+1)}\) represent the velocity and kinetic energy of the apple when leaving the plate at the time of rebound.

Pulp Firmness

After the collection of hyperspectral images, the pulp firmness of the damaged region was immediately tested by a texture analyzer (TA.XT plus, Stable Micro Systems Ltd., UK). A stainless needle probe with a diameter of \(\Phi = 2\) mm was used to penetrate the sample by a compression test at the depth of 10 mm. The speed of the probe was 1 mm/s. Three time measurements of firmness were conducted for the damaged region. The mean force of these three measurements from 2.5 to 4.5 s was calculated as the damaged pulp firmness in N. Additionally, three-time tests were also operated at the undamaged region on the opposite side along the damaged region of each apple fruit. The mean force was calculated as the undamaged pulp firmness.

Hyperspectral image acquisition and pre-processing

The NIR hyperspectral images of apples were captured by the hyperspectral imaging system in reflectance mode. The system consisted of a line-scanning spectrograph (Impsector N17, Spectral Imaging Ltd., Oulu, Finland), a charged couple device (CCD) camera (Zelos-258GV, Kappa optronics GmbH, Germany), an adjustable illumination unit consisting of four halogen tungsten lamps with a range of 0 ~ 35 W, a conveying motorized stage (PSA200-11-X, Zolix, Ltd, Beijing, China), and a computer with image acquisition software (V10E; Isuzu Optics Corp., Taiwan, China). The hyperspectral images were formed in the range of 900 to 1700 nm in this study. The spectral increment was 3.33 nm, with a total of 256 contiguous wavebands. The illumination unit was used to illuminate the sample placed on the conveying stage, and the emitted light of the sample was captured through the lens of the CCD camera to obtain image and spectral information.

To ensure the integrity of the samples’ spectra collection, the damaged area of ‘Red Fuji’ apples was placed on the conveying stage directly below the camera. The marked damaged area was perpendicular to the lens. The parameters were set as follows: The distance between the sample and the lens was 32 cm; the exposure time was 20 ms; and the moving speed of the conveying stage was 0.76 cm/s. All data were recorded by computer software. This information finally formed three-dimensional data cubes called ‘hypercubes’, which could be analyzed to ascertain the physical and chemical features of fruit simultaneously. The ‘hypercubes’ contained image information and spectral information with a dimension of \(x \times y \times \lambda\). The image space size \(x \times y\) was 320 pixel \(\times \) 194 pixel and the spectral information \(\lambda\) ranged from 900 to 1700 nm.

The hyperspectral imaging system acquired raw hyperspectral images. All raw hyperspectral images were first calibrated for reflectance due to the dark current in the CCD camera and non-uniformity of the light illumination before further data processing and analysis. Under the same condition as the acquisition of hyperspectral images, the white reference image was acquired by collecting a standard polytetrafluoroethylene (PTFE) white calibration board. A key point here was that in order to avoid spectra saturation in the corrected hyperspectral data, the white calibration board was placed at the same height as the sample. The dark reference image was acquired by turning off the lamps and covering the lens with a completely opaque camera cap. The dark and white reference images were used for calibration. The calibrated image \((I)\) was calculated by Eq. 3:

\[I = \frac{I_t - I_d}{I_w - I_d}\] (3)
where $I_r$ is the raw hyperspectral image, $I_d$ is the dark reflectance image, and $I_w$ is the white reflectance image. An ellipse region of interest (ROI) was selected in the center of the damaged area of the apple in each calibrated hyperspectral image. Thereby, the mean spectral value of ROI was calculated by ENVI 5.1 software.

Spectral Analysis

There is not a standard about which is the best type of pre-treatment for spectra. Therefore, it is necessary to pretreat spectra using different methods. Appropriate spectral pre-treatment can reduce the influence of various non-target factors on the signal and model performance. The pretreat methods included multiplicative scatter correction (MSC), standard normal variate (SNV) transformation, the first-order derivative (1st Der), and the second-order derivative (2nd Der) in this study.

Characteristic wavelength selection

Hyperspectral data have the characteristics of a high dimensionality and collinearity. Through a proper selection of the wavelength that contains the most relevant information, the computation time can be reduced, and interference from noise can be avoided. Moreover, this is an essential step for establishing a rapid detection system for real-time applications. Two methods, consisting of the SG 2nd Der and CARS method, were used for the selection of characteristic wavelengths in this study.

The SG 2nd Der is a statistical approach used to extract useful information laying in spectral data. It plays an important role in hyperspectral analysis, which can greatly reduce the correlation between variables, highlight the characteristic peak bands of the spectral curve, and obtain a new subset of variables that can explain the largest variance, thereby determining and selecting characteristic wavelengths. The subset of the characteristic peak bands of the spectral curve is selected as characteristic wavelengths. Douglas Barbin et al. used the SG 2nd Der to identify key wavelengths, and performed principal component analysis using key wavelength modeling. The results indicated that the pork category can be accurately distinguished.

The CARS method is a strategy for selecting important wavelengths based on Monte Carlo sampling and PLS regression coefficients, which employs the ‘survival of the fittest’ principle of Darwin’s evolution. The core idea is to preferentially select the variables with a large absolute value of the regression coefficient in the PLS regression process during the sampling process. Finally, a subset with the lowest root mean square error of cross-validation is considered as the optimal characteristic wavelength after a certain number of samplings. The main steps applied to select characteristic wavelengths have been described in detail.

Partial Least Squares regression

PLS regression is a powerful multi-analysis method and widely used in chemometric analysis. It was developed on the basis of multiple linear regression and principal component regression. This technique overcomes the problem of multlinearity among variables and makes the latent variable the strongest correlation with sample characteristics.

PLS regression was applied to establish independent models for each measured parameter. Five different damage degrees were used in this study. Thirty samples were randomly selected from each group, and a total of 150 samples were selected as the calibration set. For the remaining 10 samples of each group, a total of 50 samples were used as the prediction set. The quantitative relationship of PLS is built by following:

$$Y_{n \times 1} = X_{n \times k} \times B_{k \times 1} + E_{n \times 1}$$

(4)

where $X$ is the independent variable – the extracted spectral matrix with 220 bands for 150 calibration samples, $Y$ is dependent variable – the parameter for 150 calibration samples. $B$ is the matrix of regression coefficients, and $E$ is the regression residual matrix. Separate regression models of $X$ and $Y$ were established for the damaged pulp firmness, damaged area, absorbed energy, average pressure, and contact load, respectively. In this study, full cross-validation was used to calculate the sum of squared prediction residuals and optimize the calibration model. Only one sample was preserved at a time during the calculation and all other samples were used for calibration. This step continues until
each sample is treated as validation data. The performances of the developed calibration models were further validated to predict the parameters of the samples in an independent prediction set. The parameters of samples from the prediction set were predicted by the established models. The performance of the model was evaluated according to the coefficient of determination of calibration ($R_C^2$), validation ($R_{CV}^2$), and prediction ($R_P^2$), and the root-mean-square errors of the calibration (RMSEC), validation (RMSECV), and prediction (RMSEP). Generally, good models should exhibit lower RMSEC, RMSECV, and RMSEP values and higher $R_C^2$, $R_{CV}^2$, and $R_P^2$ values.

Results and discussion

Spectral characteristics

Every pixel point in the hyperspectral image cube has spectral information. The spherical surface of apples leads to a non-uniform lighting intensity distribution. The distribution of the illumination intensity is lower at the edges and higher in the portion near the center. In order to reduce the influence of the spherical surface, the average spectra of the ROI were extracted from the center of each sample. The spectra outside the 23–242 bands had a significantly lower signal-to-noise ratio because of the lower quantum efficiency and dark current of the CCD detector at the edge. In this case, only 220 bands between 23 and 242 (corresponding wavelengths from 957 to 1687 nm) were used in the subsequent data analysis process.

The raw reflectance spectra of all samples are shown in Figure 2(a). To show the sample spectra at different drop heights more clearly, the average spectra from each group are presented in Figure 2(b). As can be seen from Figure 2(b), the average spectral curves of the intact samples and bruised samples display a consistent trend. The general trends of all samples are similar. However, there are distinct discrepancies among different drop heights. The average spectra of damaged apples exhibit lower reflectance than intact apples. The average spectral reflectance of the damaged area at different heights is significantly lower than that of the intact surface. This is consistent with the conclusion that the water content in the damaged area is generally higher than that of other normal tissues. The reflectance of the samples gradually decreases with increasing drop heights. The difference between the intact and damaged spectra is mainly related to changes in apple quality.

Specifically, mechanical damage can cause tissue damage, texture deterioration, and cell rupture. Enzymes are released to cause the oxidative browning after cell rupture. Due to the physical and physiological reasons mentioned above, the moisture content increases after injury. Under the impact condition, the destroyed cell wall and chemical changes cause the light scatter in the damaged fruit tissue to change, which leads to the reflectivity of bruised fruit displaying differences. Besides, the degree of impact damage is closely correlating to the change of spectral reflectance, which inspired us

![Figure 2](image-url)  
Figure 2. (a) Raw reflectance spectra and (b) average reflectance spectra of intact and bruised regions on apples.
to establish the model to connect spectral data and damage parameters. In order to achieve the prediction of damaged parameters to apples, further analysis of the spectra and parameters is necessary.

Characteristic wavelength selection

The $R_p^2$ of the PLS regression model obtained by using different pretreat methods is shown in Table 1. The $R_p^2$ values of the spectral data without pretreatment were 0.652, 0.846, 0.780, 0.816, and 0.617, corresponding to the average pressure, contact load, damaged area, absorbed energy, and damaged firmness, respectively. It was evident that the pre-processed data did not improve the predictive performance of the model. The PLS regression model results based on full wavelengths yielded slightly better results than those based on pre-processed data. Therefore, the following steps were based on the raw full wavelength spectral data.

The second-order derivative spectra, shown in Figure 3, can be used to explain the spectral information. Figure 3 revealed greater spectral details comparing to raw full spectra. The second-order derivative spectra could resolve the broad maxima which were seen in raw spectral plots into a number of sharper peaks.\(^{[27]}\) This reflected the positions of major variability in the spectral collection. The characteristic wavelengths were determined and selected by highlighting sharper peak bands in the spectral curve. Based on Figure 3, the most important wavelengths were distinguished as being 967, 1001, 1100, 1154, 1190, 1407, and 1443 nm. The wavelengths near 967 and 1443 nm could mainly be attributed to the second and first O-H stretch overtones, which was associated with the water content of apples.\(^{[28,29]}\) The valley observed at 1154 nm was probably due to the symmetrical C-H stretch.\(^{[30]}\) The valley band at approximately 1100 nm was due to the second overtone N-H stretching.\(^{[31]}\)

Table 1. The coefficient of determination of prediction ($R_p^2$) values of different spectral pretreat methods.

| Parameter  | Raw | MSC | SNV | 1st Der | 2nd Der |
|------------|-----|-----|-----|---------|---------|
| AP (MPa)   | 0.652 | 0.603 | 0.565 | 0.569 | 0.487 |
| CL (N)     | 0.846 | 0.817 | 0.811 | 0.821 | 0.819 |
| DA (mm\(^2\)) | 0.780 | 0.779 | 0.734 | 0.726 | 0.790 |
| AE (J)     | 0.816 | 0.781 | 0.792 | 0.770 | 0.774 |
| DF (N)     | 0.617 | 0.600 | 0.570 | 0.603 | 0.570 |

AP: average pressure; CL: contact load; DA: damaged area; AE: absorbed energy; and DF: damaged firmness.

Figure 3. Distribution of selected variables by the Savitzky–Golay second-order derivative (SG 2\(^{\text{nd}}\) Der).
The process of CARS was based on the raw spectral data. The process of extracting characteristic wavelengths using CARS can be specifically expressed by Figure 4, which shows the variation path of the regression coefficients of the 220 wavelength variables in Monte Carlo sampling. The downward trend of the number of variables in Figure 4 (a) was quick at first and then flattened. It can be seen from Figure 4 (b) that with the increase of the number of samplings, the RMSECV value first decreased slowly, almost tending to the level, and reached the lowest value when the number of samplings was 26. That was to say, during the 1–25 operations, some uninformative variables were eliminated. Then, the RMSECV value continued to rise, indicating that the key information was continuously excluded from the optimal subset, leading to the worse model performance. As a result, the variable subset obtained from the 26th sampling was considered to be the optimal characteristic wavelength variable subset. This subset contained the 20 selected variables of 971, 974, 1007, 1044, 1047, 1070, 1074, 1104, 1107, 1110, 1154, 1157, 1204, 1234, 1320, 1330, 1334, 1360, 1370, and 1377 nm, respectively.

Parameter-based statistical analysis

Figure 5 shows examples of the pressure-sensitive film used for apple impacts at five drop heights. The indexes 1, 2, 3, 4, and 5 represent the effective rate, measured area, average pressure, maximum pressure, and contact load, respectively. The effective rate read from the film exceeds 85%, which indicates that the scanned data were valid and could be further analyzed. The damaged areas were 623, 955, 1081, 1394, and 1503 mm² for apple impacts from drop heights of 30, 60, 90, 120, and 150 cm, respectively. Obviously, as the drop height decreases, the damaged area declines. It can be seen from the scanning results that the smallest contact pressure is closer to the edge, but the largest pressure is not at the center point. The yellow points, which represent over-saturated data, are randomly distributed at the center of the pressure surface. Lewis et al. [1] obtained different results: Ultrasonic scans showed that the highest contact pressure was at the center of the contact area, and the pressure fell away toward the edge. Figure 5 shows that apples’ impact damage contact shape has a tendency to

![Figure 4](https://example.com/figure4.png)

**Figure 4.** Changes in the number of sampled variables, the root-mean-square errors of the validation (RMSECV), and regression coefficients path with the increase in the number of samplings.
be elliptical. Therefore, it is reasonable to use the elliptical assumption of previous fruit damage studies.

Fruit firmness refers to the resistance of the fruit pulp under impact. A higher firmness for stiffer fruit means that samples are more resistant to damage. It can be clearly seen from Figure 6 (a) that with the increase of the drop height, the damaged firmness gradually declined. This was mainly because the impact force destroyed the tissue structure of the cell wall, promoting the softening of fruit and increasing the water content. In other words, the impact force hurt the apple fruit. Mathew et al. [32] obtained similar results, which indicated that an increase in the drop height of potato lead not only to a higher percentage of bruises, but to a shift in the type of bruise damage. Figure 6 (b) and (c) indicate that the damaged area, contact load, average pressure, and absorbed energy increased with the increase in the drop height.

Due to the diversity of physiological changes of fruit itself, such as irregular shapes and heterogeneity of the internal structure, the parameter changes in this study also display certain variability. With the increase in the drop height, the parameters exhibit a certain linear change. To some extent, it can be shown that it is feasible and reasonable to characterize the degree of impact damage of apples based on the data of obtained parameters.

Partial least squares regression model results based on spectra and parameters

Table 2 presents the performance of PLS regression models based on the full wavelength of raw data. The PLS regression model achieved good results. The $R_p^2$ of the calibration set shows that the model was stable, and the $R_p^2$ of the prediction set shows that the ideal prediction result was obtained. The prediction results of the PLS regression model of measured parameters based on the characteristic wavelengths are shown in Tables 3 and 4. The PLS regression model based on the SG 2nd Der for selecting the characteristic wavelengths is slightly inferior to that based on CARS. The CARS-based results are better than the full-wavelength-based results. It is shown that the selection of the characteristic wavelength method can affect the modeling accuracy. The CARS-PLS regression model composed of fewer variables achieved good prediction results. This indicates that CARS-PLS regression has the potential to use several effective wavelengths replace the full wavelengths. Since the modeling results based on CARS are optimal, the following data analysis was based on the CARS method.

Figure 7 shows the predicted values against measured values of the modeling and prediction sets for the average pressure, contact load, damaged area, absorbed energy, and damaged firmness. The x-axis represents the actual measured value and y-axis represents the predicted value. The sample points are all distributed around the regression line, which is relatively close to the regression line, indicating that a good fit was obtained. It is shown that the characteristic wavelength extracted based on the CARS method.
Figure 6. Changes in the parameters of apples against drop height for (a) undamaged firmness and damaged firmness, (b) absorbed energy and damaged area, and (c) average pressure and contact load.
basically covers the characteristic information of apples. Furthermore, it can accurately predict the damage parameters when combined with PLS regression. Table 4 indicates that the result of the prediction set is close to the result of the calibration set and validation set. On the whole, the PLS model employed for the prediction of the contact load, absorbed energy, and damaged area achieved rather good results, which proves the strong linear correlation between the spectra of apples. The $R_p^2$ and RMSEP results of the contact load, absorbed energy, and damaged area are 0.86 and 53.80 N, 0.81 and 0.24 J, and 0.83 and 116.37 mm$^2$, respectively. It is further shown that the characteristic wavelength extracted based on the
Table 2. Prediction results of partial least squares (PLS) regression models based on the full wavelength of raw data.

| Parameter | Calibration Sets | Cross Validation Sets | Prediction Sets |
|-----------|------------------|-----------------------|-----------------|
|           | $R^2_C$ | RMSEC | $R^2_V$ | RMSECV | $R^2_P$ | RMSEP |
| AP (MPa)  | 0.73   | 0.02  | 0.69   | 0.02   | 0.65   | 0.02  |
| CL (N)    | 0.85   | 50.04 | 0.84   | 60.59  | 0.85   | 56.39 |
| DA (mm²)  | 0.82   | 125.37| 0.81   | 131.18 | 0.78   | 131.07|
| AE (J)    | 0.86   | 0.23  | 0.85   | 0.24   | 0.82   | 0.24  |
| DF (N)    | 0.69   | 0.21  | 0.66   | 0.22   | 0.62   | 0.19  |

Table 3. Prediction results of PLS regression models based on the SG 2nd Der method.

| Parameter | Calibration Sets | Cross Validation Sets | Prediction Sets |
|-----------|------------------|-----------------------|-----------------|
|           | $R^2_C$ | RMSEC | $R^2_V$ | RMSECV | $R^2_P$ | RMSEP |
| AP (MPa)  | 0.63   | 0.02  | 0.62   | 0.02   | 0.51   | 0.02  |
| CL (N)    | 0.83   | 61.99 | 0.81   | 64.04  | 0.76   | 65.90 |
| DA (mm²)  | 0.83   | 121.90| 0.81   | 130.61 | 0.77   | 133.73|
| AE (J)    | 0.85   | 0.23  | 0.84   | 0.25   | 0.73   | 0.29  |
| DF (N)    | 0.69   | 0.21  | 0.66   | 0.22   | 0.62   | 0.19  |

Table 4. Prediction results of PLS regression models based on the competitive adaptive reweighted sampling (CARS) method.

| Parameter | Calibration Sets | Cross Validation Sets | Prediction Sets |
|-----------|------------------|-----------------------|-----------------|
|           | $R^2_C$ | RMSEC | $R^2_V$ | RMSECV | $R^2_P$ | RMSEP |
| AP (MPa)  | 0.67   | 0.02  | 0.64   | 0.02   | 0.66   | 0.02  |
| CL (N)    | 0.87   | 56.05 | 0.86   | 57.87  | 0.86   | 53.80 |
| DA (mm²)  | 0.85   | 114.39| 0.85   | 118.48 | 0.83   | 116.37|
| AE (J)    | 0.86   | 0.23  | 0.85   | 0.24   | 0.81   | 0.24  |
| DF (N)    | 0.66   | 0.22  | 0.63   | 0.23   | 0.64   | 0.19  |

CARS method basically covers the characteristic information of the apple, and the PLS regression can accurately predict the damage parameters. Because only 20 variables were used in this model, the speed and efficiency were significantly improved compared with the full wavelength, which is beneficial for practical applications.

The $R_p^2$ of the average pressure is 0.66, which indicates that the prediction of the average pressure achieves slightly satisfactory results. The reason for the poor prediction result may be that the average pressure of the pressure-sensitive film was calculated by dividing the contact force by the area. Due to the limitation of the selected model, the contact force outside the range cannot display accurate data. There is a certain range of errors in the calculation. The modeling result of the damaged firmness is not good because of the heterogeneity of the fruit itself. In the future, we will find better methods for obtaining the average pressure and pulp firmness.

Conclusion

A hyperspectral imaging system in the spectral range of 900–1700 nm was used to investigate the damage degree of apples caused by the impact force in this study. The mechanical parameters, such as the average pressure, contact load, damaged area, absorbed energy, and damaged firmness during the drop test, were obtained by a pressure-sensitive film technique and high-speed camera. Four types of pre-processed methods were adopted to pretreat raw spectral data. It was found that the raw spectral data without preprocessing led to a better model performance than that after spectral pre-treatments. Two characteristic wavelength schemes, consisting of the SG 2nd Der and CARS method, were used to select characteristic wavelengths. Compared with the PLS regression model based on the raw full wavelength and SG 2nd Der, the model based on CARS to extract characteristic wavelengths had a better result. Obviously, extracting the characteristic wavelengths has the advantages of speeding up the calculation speed and reducing the data
processing time. The PLS regression model was developed to compute the mechanical parameters based on the characteristic wavelengths. It can be used to quantify the damage degree and predict multiple parameters of apples. The results confirm the applicability of NIR hyperspectral imaging technology in apples for a quantitative evaluation and prediction of the mechanical parameters. This study demonstrates the potential of the NIR hyperspectral imaging technique as a highly accurate way to quantitatively predict the mechanical parameters of apples. Further investigations will be conducted to update the database, in order to establish a more stable model, as well as to improve the robustness of the model before online application.

**Author contributions**

Conceptualization, Peipei ZHANG and Huaiwen WANG; data curation, Hongwei JI; formal analysis, Xiaochuan ZHANG and Chunhua REN; funding acquisition, Huaiwen WANG; investigation, Peipei ZHANG and Yuexuan LIU; methodology, Peipei ZHANG and Xiaochuan ZHANG; project administration, Huaiwen WANG; resources, Hongwei JI; software, Yuexuan LIU; supervision, Hongwei JI; validation, Peipei ZHANG; writing—original draft, Peipei ZHANG; writing—review and editing, Huaiwen WANG.

**Conflict of Interest**

All authors of the present paper declare that they have no conflicts of interest.

**Informed consent**

Informed consent was obtained from all individual participants included in the study.

**Funding**

This work was supported by the National Natural Science Foundation of China [No. 11772225] and Tianjin Natural Science Foundation [18JCYBJC6200].

**References**

[1] Lewis, R.; Yoxall, A.; Marshall, M. B.; Canty, L. A. Characterising Pressure and Bruising in Apple Fruit. *Wear* 2008, 264(1–2), 37–46. DOI: 10.1016/j.wear.2007.01.038.

[2] Xing, J.; Van Linden, V.; Vanzeebroeck, M.; De Baerdemaecker, J. Bruise Detection on Jonagold Apples by Visible and Near-infrared Spectroscopy. *Food Control*. 2005, 16(4), 357–361. DOI: 10.1016/j.foodcont.2004.03.016.

[3] Kittawee, U.; Pathaveerat, S.; Srirunguang, T.; Slaughter, D. Mechanical Bruising of Young Coconuts. *Biosyst. Eng*. 2011, 109(3), 211–219. DOI: 10.1016/j.biosystemseng.2011.04.002.

[4] ElMasry, G.; Wang, N.; Vigneault, C.; Qiao, J.; ElSayed, A. Early Detection of Apple Bruises on Different Background Colors Using Hyperspectral Imaging. *LWT Food Sci. Technol.* 2008, 41(2), 337–345. DOI: 10.1016/j.lwt.2007.02.022.

[5] Hernández-Sánchez, N.; Hills, B. P.; Barreiro, P.; Marigheto, N. An NMR Study on Internal Browning in Pears. *Postharvest Biol. Technol.* 2007, 44(3), 260–270. DOI: 10.1016/j.postharvbio.2007.01.002.

[6] Diels, E.; Van Dael, M.; Keresztes, J.; Vanmaercke, S.; Verboven, P.; Nicolai, B.; Saey, W.; Ramon, H.; Smeets, B. Assessment of Bruise Volumes in Apples Using X-ray Computed Tomography. *Postharvest Biol. Technol.* 2017, 128, 24–32. DOI: 10.1016/j.postharvbio.2017.01.013.

[7] Azadbakht, M.; Vahedi Torshiz, M.; Mahmoodi, M. J. The Use of CT Scan Imaging Technique to Determine Pear Bruise Level Due to External Loads. *Food Sci. Nutr.* 2019, 7(1), 273–280. DOI: 10.1002/fsn3.882.

[8] Zhou, Y.; Mao, J.; Liu, T.; Zhou, W.; Chen, Z. Discriminating Hidden Bruises in Loquat by Attenuation Coefficients Estimated from Optical Coherence Tomography Images. *Postharvest Biol. Technol.* 2017, 130, 1–6. DOI: 10.1016/j.postharvbio.2017.04.002.

[9] Plaza, A.; Benediktsson, J. A.; Boardman, J. W.; Brazil, J.; Bruzzone, L.; Camps-Valls, G.; Chanussot, J.; Fauvel, M.; Gamba, P.; Gualtieri, A.; et al. Recent Advances in Techniques for Hyperspectral Image Processing. *Remote Sensing of Environment*. 2009, 113, S110–S122. DOI: 10.1016/j.rse.2007.07.028.
[10] Lorente, D.; Aleixos, N.; Gómez-Sanchis, J.; Cubero, S.; Garcia-Navarrete, O. L.; Recent Advances, B. J. Applications of Hyperspectral Imaging for Fruit and Vegetable Quality Assessment. Food Bioprocess. Technol. 2011, 5(4), 1121–1142. DOI: 10.1007/s11947-011-0725-1.

[11] Liu, Q.; Sun, K.; Peng, J.; Xing, M.; Pan, L.; Tu, K. Identification of Bruise and Fungi Contamination in Strawberries Using Hyperspectral Imaging Technology and Multivariate Analysis. Food Anal. Methods. 2018, 11(5), 1518–1527. DOI: 10.1007/s12161-017-1136-3.

[12] Pan, L.; Sun, Y.; Xiao, H.; Gu, X.; Hu, P.; Wei, Y.; Tu, K. Hyperspectral Imaging with Different Illumination Patterns for the HOLLowness Classification of White Radish. Postharvest. Biol. Technol. 2017, 126, 40–49. DOI: 10.1016/j.postharvbio.2016.12.006.

[13] Li, J.; Rao, X.; Ying, Y. Detection of Common Defects on Oranges Using Hyperspectral Reflectance Imaging. Comput. Electron. Agric. 2011, 78(1), 38–48. DOI: 10.1016/j.compag.2011.05.010.

[14] Cen, H.; Lu, R.; Zhu, Q.; Mendoza, F. Nondestructive Detection of Chilling Injury in Cucumber Fruit Using Hyperspectral Imaging with Feature Selection and Supervised Classification. Postharvest. Biol. Technol. 2016, 111, 352–361. DOI: 10.1016/j.postharvbio.2015.09.027.

[15] ElMasry, G.; Wang, N.; Vigneault, C. Detecting Chilling Injury in Red Delicious Apple Using Hyperspectral Imaging and Neural Networks. Postharvest. Biol. Technol. 2009, 52(1), 1–8. DOI: 10.1016/j. postharvbio.2008.11.008.

[16] Zhang, B.; Fan, S.; Li, J.; Huang, W.; Zhao, C.; Qian, M.; Zheng, L. Detection of Early Rottenness on Apples by Using Hyperspectral Imaging Combined with Spectral Analysis and Image Processing. Food Anal. Methods. 2015, 8, 2075–2086. DOI: 10.1007/s12161-015-0097-7.

[17] Fan, S.; Li, C.; Huang, W.; Chen, L. Detection of Blueberry Internal Bruising over Time Using NIR Hyperspectral Reflectance Imaging with Optimum Wavelengths. Postharvest. Biol. Technol. 2017, 134, 55–66. DOI: 10.1016/j. postharvbio.2017.08.012.

[18] Wang, H.; Marker Identification Technique for Deformation Measurement. Adv. Mech. Eng. 2013, 5, 246318. DOI: 10.1155/2013/246318.

[19] Zhang, X.; Chu, X.; Ji, H.; Wang, Y. Effect of Freezing Rate on the Onion Cell Deformation Evaluated by Digital Image Correlation. Food Anal. Methods. 2016, 9(11), 3125–3132. DOI: 10.1007/s12161-016-0504-8.

[20] Pan, B.; Qian, K.; Xie, H.; Asundi, A. TOPICAL REVIEW: Two-dimensional Digital Image Correlation for In-plane Displacement and Strain Measurement: A Review. Measurement Ence & Technology. 2009, 20, 152–154. DOI: 10.1088/0022-0466/6/2001.

[21] Rinnan, Å.; Berg, F. V. D.; Engelsen, S. B. Review of the Most Common Pre-processing Techniques for Near-infrared Spectra. TrAC Trends Anal. Chem. 2009, 28(10), 1201–1222. DOI: 10.1016/j.trac.2009.07.007.

[22] Barbín D, Elmasry G, Sun D W, et al. Near-infrared hyperspectral imaging for grading and classification of pork [J]. Meat Science, 2012, 90(1):259–268.

[23] Li, H.; Liang, Y.; Xu, Q.; Cao, D. Key Wavelengths Screening Using Competitive Adaptive Reweighted Sampling Method for Multivariate Calibration. Anal. Chim. Acta. 2009, 648, 77–84. DOI: 10.1016/j.aca.2009.06.046.

[24] Tang, G.; Huang, Y.; Tian, K.; Song, X.; Yan, H.; Hu, J.; Xiong, Y.; Min, S. A New Spectral Variable Selection Pattern Using Competitive Adaptive Reweighted Sampling Combined with Successive Projections Algorithm. Analyst. 2014, 139, 4894–4902. DOI: 10.1039/c4an00837c.

[25] Geladi, P.; Kowalski, B. R. Partial Least-squares Regression: A Tutorial. Anal. Chim. Acta. 1986, 185, 1–17. DOI: 10.1016/0003-2670(86)80028-9.

[26] Lu, R.; Chen, H.; Huang, M.; Ariana, D. P. Spectral Absorption and Scattering Properties of Normal and Bruised Apple Tissue. Trans. ASABE. 2010, 53(1), 263–269. DOI: 10.1033/2013.29491.

[27] Esquerre, C.; Gowen, A. A.; O’Donnell, C. P.; Downey, G. Initial Studies on the Quantitation of Bruise Damage and Freshness in Mushrooms Using Visible-Near-Infrared Spectroscopy. J. Agric Food Chem. 1993-1907, 2009 (57). DOI: 10.1021/jf080309c.

[28] Bünning-Pfaue, H.; Analysis of Water in Food by near Infrared Spectroscopy. Food Chem. 2003, 82(1), 107–115. DOI: 10.1016/S0308-8146(02)00583-6.

[29] Howard, D. L.; Kjaergaard, H. G. Influence of Intramolecular Hydrogen Bond Strength on OH-stretching Overtones. J. Phys. Chem. A. 2006, 110, 10245–10250. DOI: 10.1021/jp063631.

[30] Ren, Y.; Shimoyama, M.; Ninomiya, T.; Matsukawa, K.; Inoue, H.; Noda, I.; Ozaki, Y. Two-Dimensional Near-Infrared Correlation Spectroscopy Studies on Composition-Dependent Spectral Variations in Ethylene/Vinyl Acetate Copolymers: Assignments of Bands Due to Ethylene Units in Amorphous, Disordered, and Orthorhombic Crystalline Phases. Evolution; International Journal of Organic Evolution. 1999, 53(3), 919–926. DOI: 10.1006/jores.1999.1021.

[31] Rvasco, B.; Miller, C.; King, T. Utilization of NIR Spectroscopy to Estimate the Proximate Composition of Trout Muscle with Minimal Sample Pretreatment. J. Agric. Food Chem. 1991, 39(1), 67–72. DOI: 10.1021/jf00001a012.

[32] Mathew, R.; Hyde, G. Potato Impact Damage Thresholds. Trans. ASAE. 1997, 40(3), 705–709. DOI: 10.13031/2013.21290.