Bulgarian Chemical Communications, Volume 56, Special Issue D (pp. 106-115) 2024 DOI: 10.34049/bcc.56.D.S1P38

Classification of apple varieties by VIS-NIR diffuse reflectance spectroscopy and chemometrics

J. Cruz[1], G. Gergov[2*], M. Tarapoulouzi[3], E. Kirilova[2], O. Kostadinova[4], K. Nikolova[5]

[1]*Escola Universitària Salesiana de Sarrià Passeig Sant Joan Bosco 74 08017 Barcelona, Catalonia, Spain*
[2]*Bulgarian Academy of Sciences, Institute of Chemical Engineering, Acad. Georgi Bontchev Str., Bl.103, Sofia 1113, Bulgaria*
[3]*Department of Chemistry, Faculty of Pure and Applied Science, University of Cyprus, P.O. Box 20537, Nicosia CY-1678, Cyprus*
[4]*Bulgarian Academy of Sciences, Institute of Electrochemistry and Energy Systems "Acad. E. Budevski", Acad. Georgi Bonchev Str. Bl.10, 1113 Sofia, Bulgaria*
[5]*Department of Physics and Biophysics, Medical University of Varna, Faculty of Pharmacy, 84 Tzar Osvoboditel Blvd, 9002 Varna, Bulgaria*

Received: November 3, 2023; Revised: April 11, 2024

The new promising VIS NIR-ES (visible near infrared extra sensitive) spectrometer was used to identify a variety of apples. The aim of this study was to investigate the applicability of VIS-NIR spectroscopy combined with chemometric methods for the classification of different varieties of apples. The study was performed by analyzing the diffuse reflectance spectra of three different types of Chinese apples: the Fuji apple, the Red Star apple, and the Gala apple. To achieve this, after suitable preprocessing, a variable selection algorithm called interval partial least squares discriminant analysis (iPLS-DA) was used to identify the most significant spectral range. This range was then compared with the results obtained by PLS-DA (using the entire NIR spectrum) to discriminate apples of different species. The obtained results have shown that the iPLS-DA model outperforms the PLS-DA application in the entire spectral range when it comes to apple cultivar identification. The best model achieved an impressive 97.77% accuracy in the calibration set and 100% accuracy in the prediction set. These results show that the use of NIR spectroscopy is a capable method for identifying the apple variety and even the growing region.

Keywords: VIS-NIR, PLS-DA, iPLS-DA, apple varieties, classification

INTRODUCTION

There are a large number of apple cultivars available around the world. Several cultivars dominate the global supply and production: 'Fuji', 'Red star', 'Royal Gala', 'Golden Delicious', 'Granny Smith'. However, cases of mixing different varieties of fruit at harvest or in the market are observed. Therefore, incorrectly labeled or manipulated apples have become more frequent. The latter is becoming a global problem whose economic and social impact is difficult to evaluate. This fact highlights the need to implement increasingly accessible methods of detection and authentication of foods that conform to what is declared.

Numerous studies have demonstrated the effectiveness of VIS-NIR and NIR techniques in assessing apple classification and quality, showing significant potential for the food industry according to the quality of apples without damaging them [1]. Another challenge is the development of affordable and portable spectroscopic devices that can be easily integrated into fruit packaging lines for quick and real-time quality assessment, making these technologies more widely applicable.

The study of Pissard *et al.* [2] confirmed the importance of NIR spectroscopy to determine phenolic compounds and dry matter in apple peel and flesh separately to evaluate fruit quality. It was proved that different apple cultivars have different content of those parameters between the peel and pulp. The outcome of their study is in agreement with our findings which shows the differentiation of apple varieties based on NIR spectroscopy. Similarly, Beghi *et al*. [3] have used a portable VIS-NIR system in combination with PLSR to predict the total phenolic content in two apple varieties, "Stark Red Delicious" and "Golden Delicious". The latter variety had low phenolic concentration, and classification was successful. Pissard *et al.* [4] evaluated the performance of benchtop and portable devices (MicroNIR) utilizing NIR spectroscopy to explore the possibility of assessing various quality parameters in apples, including soluble solids content, titratable acidity, pulp firmness, and starch-iodine index. The results, based on Partial Least Squares (PLS) models, showed that the coefficient of determination (R^2) and the root mean square error of cross-validation (RMSECV) values were quite

* To whom all correspondence should be sent:
E-mail: *ggergov187@gmail.com*

similar for both devices. This suggests that MicroNIR provides performance comparable to that of the XDS device. The best results were achieved with the Least Squares Support Vector Machines (LS-SVM) chemometric method. The outcome supports the idea of our findings in terms of promoting NIR-based technologies and chemometrics in apple classification. Li *et al*. [5] studied combining NIR spectra with PCA, successive projections algorithm (SPA) in apple analysis. Three different pattern recognition methods, namely, the backpropagation neural network (BPNN), SVM, and extreme learning machine (ELM) were also applied to create models for distinguishing apples based on their varieties and geographical origins. Notably, the SPA-ELM model achieved an impressive 98.33% accuracy in identifying apples in the calibration set and 96.67% accuracy in the prediction set. This research suggests that NIR spectroscopy is a viable approach for identifying the variety and cultivation region of apple samples.

Moreover, Xu *et al*. [6] recently differentiated different varieties of apples by combining a similarity-based particle swarm optimization algorithm with the possibilistic fuzzy c-means (PFCM) algorithm, thus SPSO-PFCM. In addition, MSC and PCA chemometric methods eliminated the interference and reduced the complexity of the spectral data. Their results convincingly demonstrate that combining NIR diffuse reflectance with SPSO-PFCM clustering is an effective method for classifying different apple varieties. Cortés *et al.* [7] assessed five different apple varieties by using in-line VIS-NIR reflectance spectroscopy. To extract the most critical information from the spectra, PCA was employed. Seven principal components were then used in LDA and quadratic discriminant analysis (QDA). The results revealed that QDA was the most effective in-line classification method, achieving success rates of 98% for red apple varieties and 85% for yellow apple varieties. This study confirms that the in-line application of VIS-NIR spectroscopy is potentially feasible for accurately detecting apple varieties. Ongoing research and development in spectroscopic techniques for apple quality assessment can potentially enhance the efficiency, quality, and safety of the apple supply chain.

The aim of the present study was to evaluate the applicability of VIS-NIR spectroscopy combined with preprocessing and chemometric tools to classify three different types of Chinese apples: the Fuji apple, the Red Star apple, the Gala apple, and compare the results with the ones from Li *et al*. article [5] where data from this study have been obtained. To achieve this, a variable selection algorithm called interval partial least squares discriminant analysis (iPLS-DA), after applying appropriate preprocessing, was used to identify the best spectral range. This range was then compared with the results obtained by applying classical PLS-DA using the entire NIR spectrum to discriminate apples of different varieties.

MATERIALS AND METHODS

Sample Preparation Procedure

For the purpose of this research, data from the article from Li *et al.* [6] were used. A total of 300 apples were selected to ensure representation and accuracy. Among these, 100 Fuji apples, 100 Red Star apples, and 100 Gala apples were used. Apples were selected from two prominent local markets. All selected apples exhibited smooth and unblemished skin, ensuring the highest quality for our study.

Before conducting any measurements, strict protocols were followed. First, all apple samples were placed in airtight polyethylene bags and stored in a refrigerator, maintaining a consistent cold temperature of 4 ± 1°C for a duration of 2 days. Subsequently, after the designated storage period, the apples were removed from the refrigerator, thoroughly washed with clean water, meticulously wiped dry, and then left to acclimatize at room temperature (24 ± 2°C) for approximately 3 hours. Only after these meticulous preparation steps the apples were suitable for spectral measurements.

Spectra Acquisition Procedure

An Ocean Optics USB2000-VIS-NIR-ES spectrometer was employed, sourced from Ocean Optics in the USA, equipped with HL-2000 tungsten halogen light sources also from Ocean Optics, and optical fiber reflection probes (QR600-7-VIS-NIR, Ocean Optics, USA).

NIR diffuse reflectance spectra were captured within the wavelength range of 400 to 1021nm, with an interval of approximately 0.33 nm, resulting in a total of 1888 variables for each spectrum. The data collection and transformation of spectra were facilitated using Ocean View software (Ocean Optics, USA). All measurements took place under controlled room temperature conditions, precisely at 24 ± 2°C.

Before the commencement of spectral measurements, a critical step involved the spectrometer being powered on for a minimum of 1 hour to stabilize and reach an optimal operating temperature. This ensured that the instrument

performed consistently throughout the data acquisition process.

During the actual spectral measurements, close proximity of the NIR optical fiber probe to the surface of the apple samples was maintained. This approach was adopted to minimize surface reflectance and eliminate any potential interference from the surrounding air.

A systematic approach was implemented to ensure robust data representation and reliability. For each intact apple sample, diffuse reflectance spectra were obtained at 15 distinct points, which were randomly selected along the equator of the apple. At each of these points, the spectral scan was repeated 10 times, resulting in a total of 150 scans. These 150 scans were then averaged to form a comprehensive and representative spectral dataset for each apple, subsequently serving as the dataset for chemometric analysis.

Data Analysis

Principal component analysis (PCA) and partial least square discriminant analysis (PLS-DA) models were applied by using Solo, version 9.2.1, a software solution developed by Eigenvector Research, Inc., headquartered in Wenatchee, WA, USA. Diverse spectral ranges were systematically explored by applying data preprocessing techniques to get the most suitable prediction models.

Principal Component Analysis (PCA)

Principal Component Analysis, often abbreviated as PCA, stands as a highly effective data mining technique that has found widespread application in spectral data analysis. The core principle behind PCA revolves around dimensionality reduction and orthogonalization of the original multidimensional dataset. The ultimate aim is to derive a set of linearly uncorrelated variables, termed principal components (PCs), with several key objectives in mind.

One primary goal is to minimize the risk of overfitting, a phenomenon where a model becomes too complex and starts fitting noise in the data rather than the true underlying patterns. Additionally, PCA enhances the training procedure's computational efficiency, making it more manageable and resource-efficient.

These principal components are calculated simultaneously through a single matrix decomposition, which extracts essential information from the original data while significantly reducing the number of variables. The first principal component is designed to capture as much variability as possible in the original dataset, thus providing a comprehensive overview of the data's main patterns. Subsequent components follow, each exhibiting lower variance than its predecessors.

One of the key strengths of PCA is its ability to transform high-dimensional data into a lower-dimensional representation while preserving the essential information present in the original dataset. Despite this reduction in variable numbers, the principal components remain powerful descriptors that effectively encapsulate the majority of the original data's variance. This capacity to condense information while retaining its significance is a hallmark of PCA's utility in data analysis and dimensionality reduction [8].

Partial Least Squares Discriminant Analysis (PLS-DA). Partial Least Squares Discriminant Analysis, commonly called PLS-DA, is a superVISed classification algorithm that leverages the principles of PLS regression and linear discriminant analysis to effectively categorize datasets into distinct classes. This powerful technique establishes a crucial connection between predictor variables and response variables by employing a reduced number of latent variables. The primary aim of PLS-DA is to maximize the covariance between predictor variables and response variables, thus unraveling the underlying relationships within the data.

In the realm of binary classification, the PLS1 variant is frequently utilized. Here, the response variable typically assumes values of 0 or 1, signifying whether a data point belongs to a specific class or not.

iPLS-DA Classification. Interval PLS (iPLS) is a data modeling technique designed to enhance prediction accuracy by selecting a subset of variables from a dataset and optimizing performance using all available variables. iPLS employs a sequential and exhaustive search approach to identify the best individual or combination of variables for the task. In the context of iPLS, an "interval" can refer to either a single variable or a "window" encompassing adjacent variables. This concept of an "interval" is particularly relevant in situations where adjacent variables are interrelated, such as in spectroscopically correlated or time-correlated datasets, where variables in proximity exhibit related behaviors. For the discussion, an "interval" will be selected while recognizing that it may encompass one or more variables. The iPLS process initiates by creating individual Partial Least Squares (PLS) models, each utilizing only one of the predefined variable intervals. For instance, if a dataset has 100 defined intervals, the initial step involves calculating 100 models, each corresponding to a distinct interval. Cross-validation is conducted for each of

these models, and the interval yielding the lowest model root-mean-square error of cross-validation (RMSECV) is chosen as the optimal single-interval model, denoted as I1. If the objective is to select just one interval, the algorithm can conclude at this point, providing the selected interval as the outcome. However, if the aim is to incorporate multiple intervals, thereby enriching the information available to the model and potentially improving its performance, additional cycles are executed. In the subsequent cycle, the first selected interval (I1) is retained in all models and combined sequentially with each of the other remaining intervals. This process generates a new set of PLS models for each combination. Once again, using RMSECV as the guiding metric, the best combination of two intervals (I1 and an additional interval, denoted as I2) is determined. It's important to note that, at this stage, the first selected interval (I1) remains fixed. This procedure is iteratively repeated for as many intervals as needed (up to In), enhancing the model's capacity to harness information and optimize predictive performance [9].

As for the full range model, our approach started by dividing our sample dataset into two subsets, after removing one sample as outlier: a training set comprising 225 samples, meticulously selected using the Kennard & Stone algorithm [10], and a prediction set consisting of 74 samples This partitioning facilitated our models' development and subsequent evaluation, a critical step in assessing their applicability to external samples. In line with the previous methodology, we pursued developing and evaluating various chemometric models, mirroring the procedures employed in the full-range model. These models were fine-tuned to optimize performance.

Spectral Data Treatment and Model Construction

To harness the full potential of our spectral data, a comprehensive journey of data treatment and model development was performed. This encompassed the utilization of Solo, version 9.2.1, developed by Eigenvector Research, Inc. in Wenatchee, WA, USA.

Various spectral ranges and data preprocessing techniques underwent a thorough evaluation to ascertain the creation of robust prediction models.

Classification models were meticulously crafted using a split of 2/3 of the samples for calibration and 1/3 for validation, ensuring a representative dataset that included samples from all batches. Sample selection was executed with precision, employing the Kennard & Stone algorithm [9].

The exploration delved into multiple data preprocessing techniques, including Standard Normal Variate (SNV), Mean Center (MC), first and second derivatives of Savitzky–Golay, employing various window sizes and polynomial orders, Multiplicative Scattering Correction (MSC), Orthogonal Signal Correction (OSC), and combinations thereof.

In the final stages, classification models were meticulously constructed through Partial Least Squares Discriminant Analysis (PLS-DA). PLS-DA, a linear classification method rooted in the PLS regression algorithm [11], was employed for its proven effectiveness.

Performance evaluation was a paramount aspect of our methodology. This was achieved through the analysis of contingency tables and the calculation of sensitivity, specificity, and precision. Sensitivity, as defined in Equation (1), measured the model's proficiency in correctly classifying positive samples as positive:

$$\text{Sensitivity} = TP / (TP + FN) \quad (1)$$

where: TP = number of positive samples correctly classified as positive; FN = number of positive samples erroneously classified as negative.

Specificity, captured by Equation (2), assessed the model's ability to accurately classify negative samples as negative:

$$\text{Specificity} = TN / (TN + FP) \quad (2)$$

where: TN = number of negative samples correctly classified as negative; FP = number of negative samples erroneously classified as positive.

Precision, as per Equation (3), quantified the proportion of all samples classified as one class that genuinely belonged to that class:

$$\text{Precision} = TP / (TP + FP) \quad (3)$$

The constructed PLS-DA models underwent rigorous internal validation through the venetian blinds cross-validation method. Key model parameters, including the R^2 model, root mean squared error of calibration (RMSEC) (as defined in Equation 4), and root mean squared error of cross-validation (RMSECV) (as defined in Equation 5), were scrutinized for the selection of the optimal model.

Further evaluation of the model's predictive capacity was executed using distinct statistical parameters, such as the root mean squared error of prediction (RMSEP), applied to the validation set samples (as defined in Equation 6). This thorough process ensured the robustness and accuracy of our predictive and classification models.

$$RMSEC = \sqrt{\frac{\sum_{i=1}^{n})y_i-\hat{y}_i(^2}{n}} \quad (4)$$

$$RMSECV = \sqrt{\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{n}} \quad (5)$$

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{n}} \quad (6)$$

In the context of our analysis, where 'yi' represents the reference value for validation set sample 'i', 'ŷi' signifies the predicted value for the same validation set sample 'i', and 'n' stands for the total number of samples within the validation set, a critical step was taken to optimize our model.

A plot illustrating the explained variance against the number of factors was generated to select the optimal number of factors. This VISual representation played a pivotal role in guiding our selection process.

Subsequently, an initial model was honed by choosing the factors that yielded the lowest Root Mean Squared Error of Prediction (RMSEP) for the validation set. Furthermore, our selection criteria incorporated a preference for factors that exhibited sensitivity and specificity coefficients that closely approached a value of 1. This comprehensive approach ensured the refinement of the final model to enhance its predictive accuracy and reliability.

RESULTS AND DISCUSSION

Vis-NIR spectra

As mentioned in the introduction and materials and methods section, Vis-NIR spectra from Li *et al.* [6] have been used for the purpose of classifying three different types of apples: the Fuji apple, the Red Star apple, and the Gala apple.

Figure 1 shows the mean apple spectra from the 3 apple classes. As it can be observed, Fuji and Red Star present similar spectra while Golden Gala, has a slight difference due to its different color. So that by observing the spectra the principal challenge from this study would be the differentiation of Red Star and Golden Gala varieties.

Classification models. With this data, non-supervised PCA models and supervised PLS-DA models are performed, as shown in the following sections.

PLS-DA classification using the full spectral range. Our dataset was thoughtfully divided into two distinct subsets: a training set consisting of 225 (2/3) samples for the development of our predictive models and a prediction set comprising 74 (1/3) samples. This segregation was achieved utilizing the Kennard & Stone algorithm. Various chemometric models were meticulously formulated and evaluated to enhance the model's efficacy and applicability. These models were tailored using carefully selected data treatment techniques.

The array of data treatments encompassed fundamental methods such as mean centering (MC), auto-scaling, and advanced techniques, including Savitzky-Golay first (FD) and second derivative (SD), multiplicative scattering correction (MSC), standard normal variation (SNV), orthogonal signal correction (OSC), and inventive combinations thereof.

Our approach commenced with an unsupervised classification analysis, primarily employing Principal Component Analysis (PCA) to take a general overview of the samples by using the Near-Infrared (NIR) spectra. The first two principal components in the PCA model encapsulated an impressive 92.30% of the total variation. However, it was evident from Figure 2 that the unsuperVISed model revealed a predominant cluster, indicating a lack of distinct clusters for the three apple categories when the full 400 to 1021nm spectral range was used.

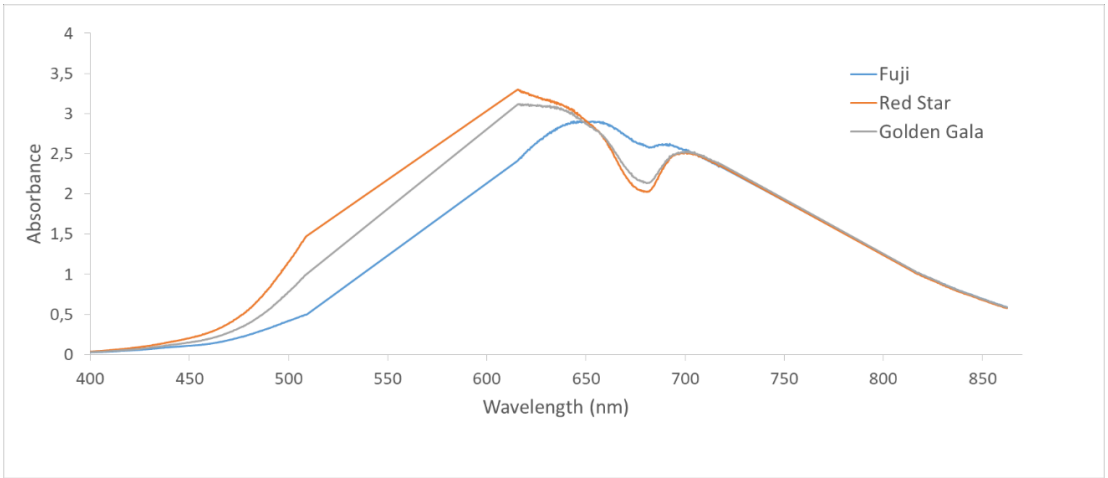

Figure 1. Mean Visible-NIR spectra from Fuji apple, Red Star apple, and Golden Gala apple.

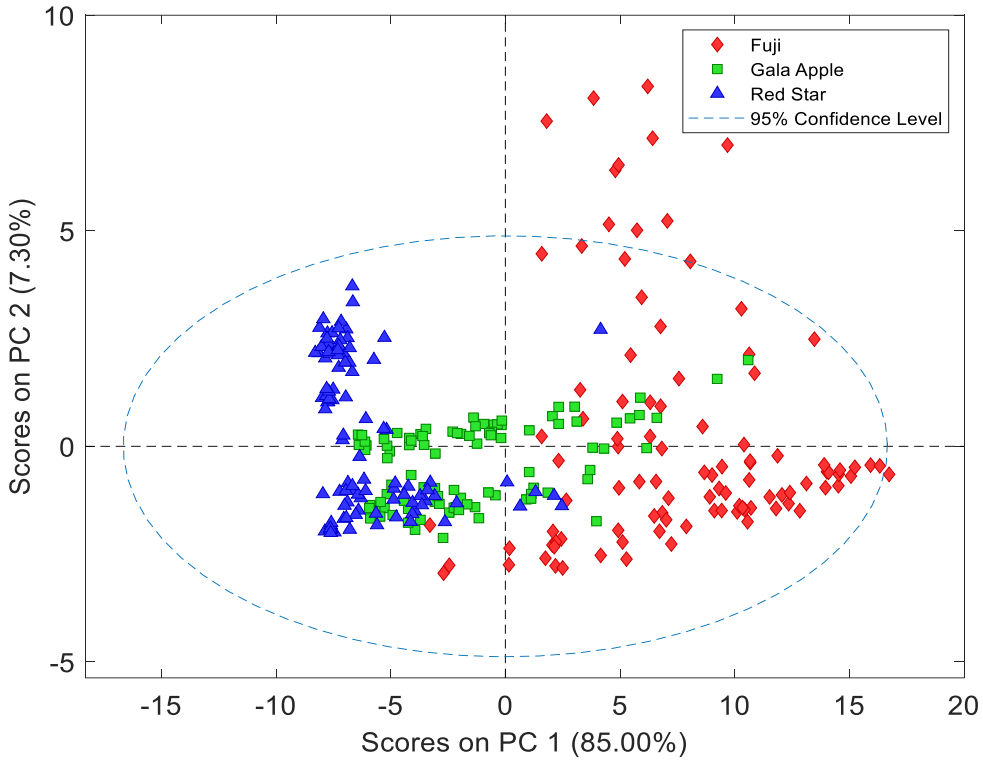

Figure 2. PCA score plot PC1 *vs*. PC2 of all apple SNV pretreated full VIS-NIR spectra data.

Table 1. Key model parameter from the SNV full range PLS-DA.

	Calibration			Cross Validation			Prediction		
	Fuji	Gala	Red Star	Fuji	Gala	Red Star	Fuji	Gala	Red Star
RMSE	0.194	0.264	0.263	0.198	0.272	0.268	0.288	0.339	0.306
R^2	0.814	0.704	0.693	0.806	0.685	0.680	0.687	0.442	0.582

Table 2. Confusion Table obtained from the SNV full range PLS-DA.

	Cross Validation			Prediction		
	Actual class			Actual class		
	Fuji	Gala	Red Star	Fuji	Gala	Red Star
Predicted as Fuji	62	4	0	37	1	0
Predicted as Gala	1	81	1	0	13	2
Predicted as Red Star	0	0	76	0	1	21
Sensitivity	1.000	0.988	0.909	0.973	1.000	0.957
Specificty	0.981	0.993	0.973	0.946	0.900	0.961

A Partial Least Squares Discriminant Analysis (PLS-DA) supervised classification model was performed to refine the classification. By utilizing predictive components, PLS-DA significantly improved class separation. The selection of the most suitable number of latent variables was determined based on the lowest value of the Root Mean Square Error of Cross-Validation (RMSECV). In our case, 6 latent variables were chosen to construct the classification model.

The performance evaluation of our classification model encompassed both internal validation (cross-validation) and external validation using the prediction set. Our model exhibited commendable R2 values for calibration, cross-validation, and prediction, as showcased in Table 1.

Furthermore, the Root Mean Square Error of Calibration (RMSEC), Root Mean Square Error of Cross-Validation (RMSECV), and Root Mean Square Error of Prediction (RMSEP) for all three apple classes demonstrated excellent agreement. This alignment indicated that the RMSECV value effectively approximated the standard error of prediction observed for the test set.

Table 2 shows the confusion table for the 3 categories with very good results. Only 2 samples for Gala apples and 2 for Redstar are misclassified in the validation set, while in the cross-validation, only 1 sample for Fuji apples, 4 for Gala apples, and 1 for Redstar are misclassified.

The ROC curves provided a comprehensive view of the models' performance. Sensitivity and specificity (Table 2) were integral components of our evaluation criteria. Sensitivity, indicative of the rate of correctly identified samples within a specified class, was plotted against 1-specificity, representing the rate of correctly identified samples within different classes.

The Area Under the Curve (AUC) of the ROC plots was instrumental in assessing method performance. The AUC value, which ranges from 0.5 (indicative of random decision) to 1 (representing a perfect model), offered insights into the model's accuracy. Remarkably, the AUC values for the full range PLS-DA model ranged from 0.975 for Gala apples to an impressive 0.996 for Fuji apples. According to the suggestion of Swets, models could be classified as non-informative (AUC=0.5), less accurate (0.5<AUC≤0.7), moderately accurate (0.7<AUC≤0.9), highly accurate (0.9<AUC<1) and perfect tests (AUC=1).

In addition to the confusion table, sensitivity and specificity values for the validation set and the cross-validation are close to 1. Table 3, with the Total positive results (TPR), false positive results (FPR), total negative results (TNR), false negative results (FNR), and % classification error (Err), shows for both prediction and cross-validation values close to 1 for TPR and TNR and 0 for FPR and FNR. Receiver-operating characteristic (ROC) curves were created for the data sets following the external prediction set to determine the performance of the chosen models, which can be seen in Figure 3.

Table 3. Total positive results (TPR), false positive results (FPR), total negative results (TNR), false negative results (FNR), and % classification error (Err) obtained from the cross-validation and prediction sets from the SNV full range PLS-DA

		N	TPR	FPR	TNR	FNR	Error
Cross Validation	Fuji	63	0.968	0.049	0.951	0.032	0.058
	Gala	85	0.894	0.029	0.971	0.106	0.058
	Red Star	77	0.974	0.007	0.993	0.026	0.013
Prediction	Fuji	37	1.000	0.027	0.973	0.000	0.014
	Gala	14	0.929	0.033	0.967	0.071	0.041
	Red Star	23	0.913	0.000	1.000	0.087	0.027

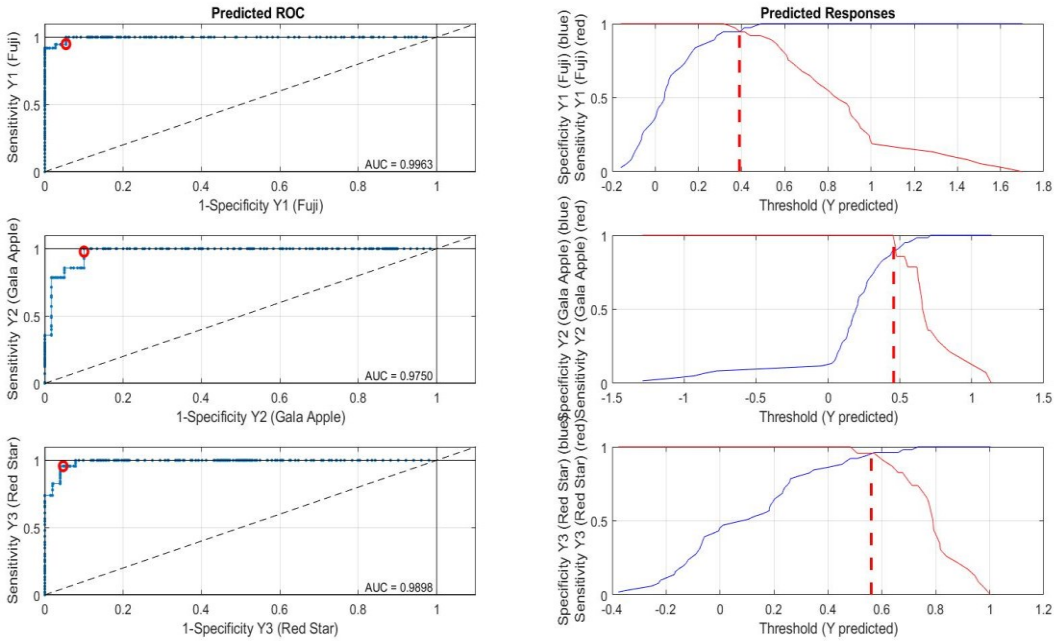

Figure 3. ROC curves obtained for full range SNV PLS-DA model

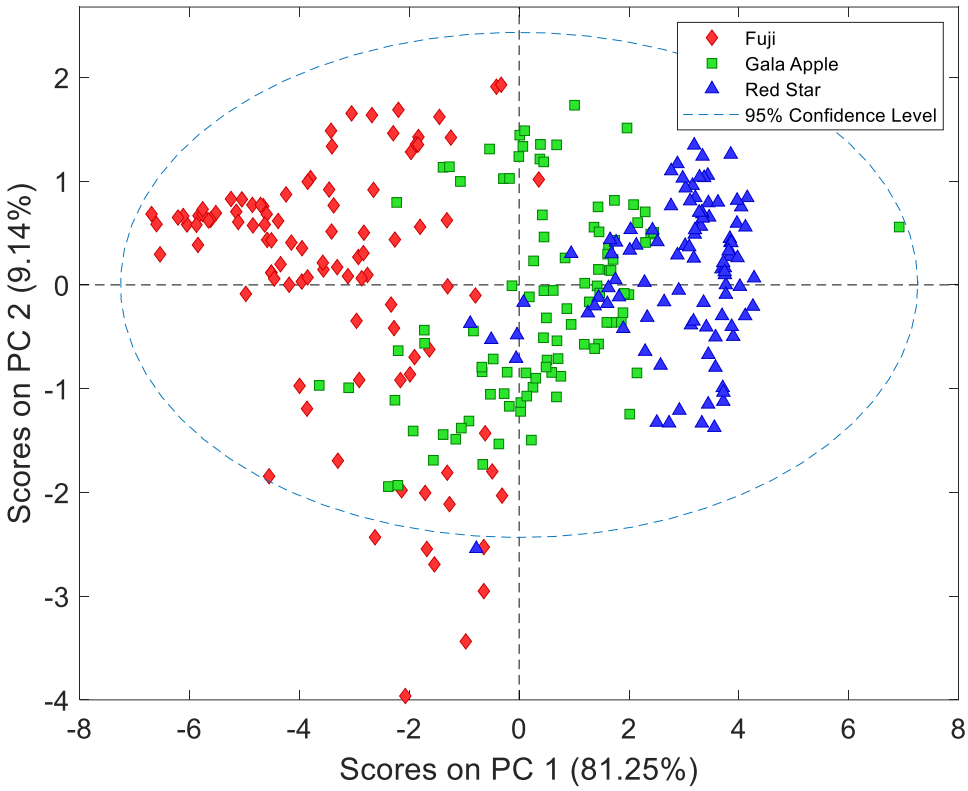

Figure 4. PCA score plot PC1 vs. PC2 of all apple SNV pretreated selected ranges of VIS-NIR spectra.

Table 4. Key model parameter from the SNV full range PLS-DA.

	Calibration			Cross Validation			Prediction		
	Fuji	Gala	Red Star	Fuji	Gala	Red Star	Fuji	Gala	Red Star
RMSE	0.166	0.275	0.183	0.195	0.323	0.202	0.267	0.412	0.269
R^2	0.864	0.679	0.851	0.812	0.556	0.818	0.730	0.276	0.663

In accordance with Swets' classification, the models were unequivocally classified as highly accurate.

iPLS-DA classification using the selected spectral range. After application of PCA method, the first two principal components (PC1 and PC2) accounted for an impressive 90.39% of the total variation. While the score plot in Figure 4 indicated the presence of a predominant cluster, a closer examination revealed that the three apple categories were distinguishable within this cluster. This observation underscored that, although PCA did not delineate specific clusters for the three apple categories, it exhibited the ability to differentiate them more effectively when the iPLS-DA 400.093-508.864, 615.223-716.993, and 815.113-861.991 nm spectral ranges were only chosen.

The implementation of the PLS-DA superVISed classification model notably enhanced the separation of distinct classes. Following a meticulous evaluation process, which considered the lowest Root Mean Square Error of Cross-Validation (RMSECV), three latent variables as the optimal number for constructing the classification model. This selection of latent variables laid the foundation for a robust classification model, and its performance was rigorously scrutinized through both internal validation (cross-validation) and external validation using the prediction set.

The results of this assessment were indicative of the model's excellence, as reflected in Table 4. Therefore, the SNV full range PLS-DA model consistently demonstrated good performance across multiple metrics, including R2 for calibration, cross-validation, and prediction. Moreover, the Root Mean Square Error of Calibration (RMSEC), Root Mean Square Error of Cross-Validation (RMSECV), and Root Mean Square Error of Prediction (RMSEP) for all three apple classes exhibited strong agreement. Table 5 presents the confusion table for the three

apple categories, yielding excellent results with no misclassifications in the validation set. In cross-validation, only 5 misclassifications were observed for the Red Star category. Supplementing the confusion table, sensitivity and specificity values for both the validation set and cross-validation were found to be closely approaching a value of 1.

Table 6 shows the model's proficiency, with values of 1 for True Positive Rate (TPR) and True Negative Rate (TNR) in the prediction set, alongside 0 values for False Positive Rate (FPR) and False Negative Rate (FNR). Cross-validation results mirror this performance, with TPR and TNR values nearing 1 and FPR and FNR registering close to 0.

To further evaluate the efficacy of our chosen models, Receiver-Operating Characteristic (ROC) curves have been created, a key step for external prediction using the prediction set. Figure 4 graphically represents the ROC curves, providing an insightful assessment of model performance. The Area Under the Curve (AUC) for the iPLS-DA range model ranged from 0.9569 for Red Star apples to 1 for Gala apples. In accordance with Swets' classification, our model for Gala apples would be classified as excellent, while the models for the other two apple varieties would be deemed highly accurate. This extensive evaluation reaffirmed the excellence and reliability of our classification model in classifying apple varieties. The classification performance of the model was also evaluated through sensitivity and specificity, as presented in Table 5.

Table 5. Confusion Table obtained from the SNV full range PLS-DA.

	Cross Validation			Prediction		
	Actual class			Actual class		
	Fuji	Gala	Red Star	Fuji	Gala	Red Star
Predicted as Fuji	63	0	3	37	0	0
Predicted as Gala	0	85	2	0	14	0
Predicted as Red Star	0	0	72	0	0	23
Sensitivity	0.984	0.882	0.987	1.000	1.000	1.000
Specificty	0.957	0.871	1.000	0.973	1.000	0.863

Table 6. Total positive results (TPR), false positive results (FPR), total negative results (TNR), false negative results (FNR), and % classification error (Err) obtained from the cross-validation and prediction sets from the SNV full range PLS-DA.

		N	TPR	FPR	TNR	FNR	Error
Cross Validation	Fuji	63	1.000	0.019	0.981	0.000	0.013
	Gala	85	1.000	0.014	0.986	0.000	0.009
	Red Star	77	0.935	0.000	1.000	0.065	0.022
Prediction	Fuji	37	1.000	0.000	1.000	0.000	0.000
	Gala	14	1.000	0.000	1.000	0.000	0.000
	Red Star	23	1.000	0.000	1.000	0.000	0.000

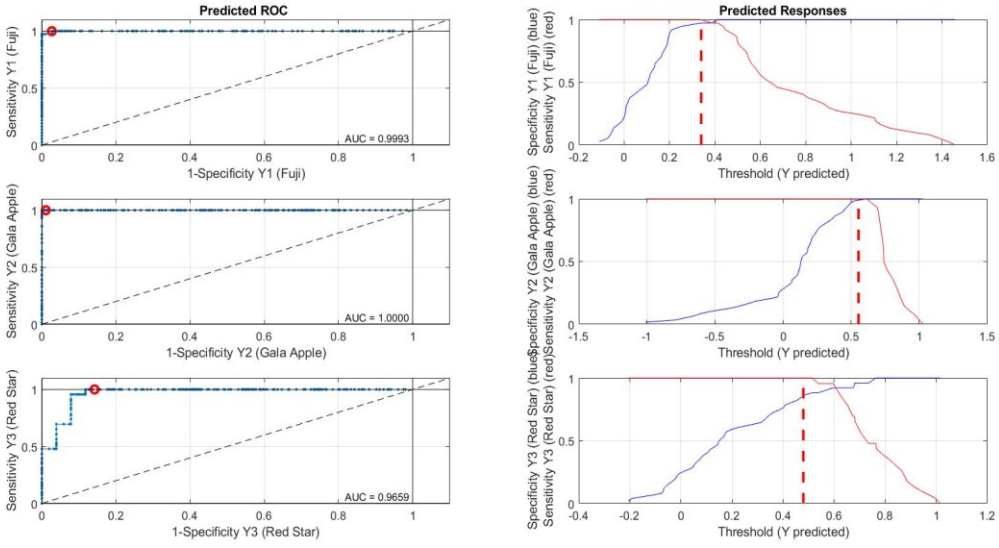

Figure 5. ROC curves obtained for iPLS-DA SNV model

Comparing the obtained results with the ones of Li *et al.* [6], we can affirm that with less complex models, PLS-DA and iPLS-DA, similar or even better accuracies are obtained than the ones obtained by SVM models and ELM models in Li *et al.* manuscript. So, models with less computation complexity can give us similar results in classifying apples among their variety.

In addition, these more complex models are more potent in classifying the apples regarding their origin. The next step in our study will be the research of the optimal conditions for obtaining comparable classification accuracies to those obtained by SVM and ELM, classifying apples regarding their origin with a more straightforward classification method.

CONCLUSIONS

In conclusion, this study highlights the efficacy of employing chemometrics in tandem with VIS-NIR direct analysis to effectively differentiate between various apple varieties, as exemplified by the successful application of PLS-DA particularly. The encouraging outcomes of the green methodology notably add another advantage to this approach.

Nevertheless, a pivotal juncture in our journey toward enhancing the classification model was the judicious selection of the most pertinent spectral range. In this crucial endeavor, the introduction of iPLS-DA emerged as a game-changer, leading to significant advancements in classifying the three distinct apple categories. This improvement was not solely confined to eliminating prediction errors within the prediction dataset; it also entailed reducing the number of latent variables within the model.

It becomes evident that the choice of an optimal spectral range not only simplifies the final model in terms of latent variables but also plays a pivotal role in elevating the predictive prowess of the model. This observation highlights the pivotal importance of spectral range optimization in refining classification models and further reaffirms its central role in improving the accuracy and efficacy of such models. Subsequently, this research demonstrates the power of precision and optimization in the field of apple spectral analysis and classification.

As mentioned in the Results section, the results obtained with PLS-DA and iPLS-DA have similar or evern better accuracies than those obtained by more complex algorithms in [6].

Acknowledgement: *This work was also supported in part by collaboration with REDTPI4.0-320RT0006 CYTED program.*

REFERENCES

1. J. Grabska, K. B. Beć, N. Ueno, C.W. Huck, *Foods*, **12**(10), 1946 (2023).
2. A. Pissard, V. Baeten, P. Dardenne, P. Dupont, M. Lateur, Use of NIR spectroscopy on fresh apples to determine the phenolic compounds and dry matter content in peel and flesh. BASE, 2018.
3. R. Beghi, A. Spinardi, L. Bodria, I. Mignani, R. Guidetti, *Food and Bioprocess Technology*, **6**, 2547 (2013).
4. A. Pissard, E. J. N. Marques, P. Dardenne, M. Lateur, C. Pasquini, M. F. Pimentel, J. A. F. Pierna, V. Baeten, *Postharvest Biology and Technology*, **172**, 111375 (2021).
5. Q. Xu, X. Wu, B. Wu, H. Zhou, *J. Food Process. Eng.*, **45**, e13993 (2022).
6. C. Li, L. Li, Y. Wu, M. Lu, Y. Yang, L. Li, *J. Spectrosc.*, 6935197 (2018).
7. V. Cortés, S. Cubero, J. Blasco, N. Aleixos, P. Talens, *Food Bioprocess. Technol.*, **12**, 1021 (2019).
8. N. Salem, S. Hussein, *Procedia Computer Science*, **163**, 292 (2019).
9. R. W. Kennard, L. A. Stone, *Technometrics*, **11**(1), 137 (1969).
10. L. Nørgaard, A. Saudland, J. Wagner, J. P. Nielsen, L. Munck, S. B. Engelsen, *Applied Spectroscopy*, **54**(3), 413 (2000).
11. S. Wold, M. Sjöström, L. Eriksson, *Chemometrics and Intelligent Laboratory Systems*, **58**(2), 109 (2001).