

Development of Method for Rapid Quantification of Glucose, Fructose and Sucrose in Mango Juice by Chemometric Techniques in De-noised FTIR Spectroscopic Data

Research Article

Uddin MN^{1*}, Majumder AK², Ahamed S¹, Saha BK³, Motalab M³

¹ BCSIR laboratories, Dhaka, Bangladesh Council of Scientific and Industrial Research (BCSIR), Dhaka, Bangladesh.

² Jahangirnagar University, Dhaka, Bangladesh.

³ Institute of Food Science and Technology (IFST), Bangladesh.

Abstract

The present study is aimed to develop a chemometrics assisted method for predicting simple sugars (glucose, fructose and sucrose) in mango juice by using the best calibration technique among Artificial Neural Network (ANN), Principal Component Regression (PCR) and Partial Least Squares Regression (PLSR) in de-noised data from Fourier Transformed Infrared (FTIR) spectrophotometer. Sixty four mixture solutions of eight different concentrations of sugars and fifteen commercial mango juices have been run in FTIR, and spectral data are used for development, validation and test of models. Standard Normal Variate (SNV), Savitzky-Golay (S-G) filtering and Multiplicative Scatter Correction (MSC) have been used for de-noising spectral data before calibration. Among the alternatives, the best prediction performance is noticed by ANN in spectral range 1500-952 cm⁻¹ and S-G filtering ($R^2 \approx 0.99$). Prediction of simple sugars concentration by ANN with FTIR spectroscopic data after de-noised with S-G filtering is a cost-effective and easy method for quantification of sugars in commercial mango juice.

Keywords: Prediction; Simple Sugars; FTIR; Spectral Data; Preprocessing; Chemometric Techniques.

Introduction

Bangladesh is one of the top ten mango growing countries in the world [1]. Mango juices are getting popularity among people of all ages. So, a number of mango juice manufacturing industries have been established. Some of the industries are exporting their product abroad. Now the quality of mango juices is a great concern both for producers and consumers. Over use of artificial sugars is one of the depressing factors and causes serious threat to public health [2]. Therefore, simple, easy and cost effective testing methods are necessary to ensure quality of mango juice for manufactures, consumers and food quality regulating authorities. Different chromatographic methods like High Performance Liquid Chromatography (HPLC), Gas Chromatography (GC), etc are used to determine concentration of sugars in foods especially fruit juice [3, 4] and honey [5]. All these methods need

tiresome and complex pretreatment of samples. Moreover, standard chemicals are used in these methods, and they require high cost for storage and disposal. On the contrary, in chemometric techniques, mathematical models are used with spectral data from instruments against known concentration of analyte where no standard chemicals are used. So, multivariate calibration methods are cheaper, faster and do not generate chemical waste [6].

Multivariate analysis techniques in Fourier Transformed Infrared (FTIR) spectroscopic data were applied to develop chemometric method for quantification of sugars in mango juice [7]. Here glucose, fructose and sucrose were estimated as a function of ripening from the model. In another study, authentication of sweeteners of mango juice using FTIR was studied by estimating Added Sugar Content (ASC), Total Sugar Solution (TSS) and Real Juice Content (RJC) [8].

*Corresponding Author:

Nashir Uddin,
BCSIR laboratories, Dhaka, Bangladesh Council of Scientific and Industrial Research (BCSIR), Dhaka-1205, Bangladesh.
Tel: 880-01912068516
Fax: 88 02 58613022
E-mail: nashir@bcsir.gov.bd

Received: November 30, 2016

Accepted: March 13, 2017

Published: March 16, 2017

Citation: Uddin MN, Majumder AK, Ahamed S, Saha BK, Motalab M (2017) Development of Method for Rapid Quantification of Glucose, Fructose and Sucrose in Mango Juice by Chemometric Techniques in De-noised FTIR Spectroscopic Data. *Int J Food Sci Nutr Diet.* 6(1), 338-344. doi: <http://dx.doi.org/10.19070/2326-3350-1700060>

Copyright: Uddin MN[©] 2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

Spectral data contain instrumental noises which to be reduced before modeling. Several de-noising techniques are used by scientists. Standard Normal Variate (SNV), Savitzky-Golay (S-G) filtering and Multiplicative Scatter Correction (MSC) are popularly used in this purpose. But their performance varies greatly according to the nature of the data. Similarly, Artificial Neural Network (ANN), Principal Component Regression (PCR) and Partial Least Squares Regression (PLSR) are most frequent chemometric techniques for calibration, and their prediction efficiencies are to be assessed to pick the best one for the spectral data used to choose the best one among the alternatives for quantification of glucose, fructose and sucrose in mango juice.

Objective of the study is to develop a rapid and inexpensive method for quantification of glucose, fructose and sucrose in commercial mango juice by using FTIR spectroscopy and comparatively better performing de-noising and calibration techniques.

Materials and Methods

Collection of Commercial Mango Juice

Fifteen commercially available mango juices of different locally manufacturing companies were collected from different outlets. At first, concentration of glucose, fructose and sucrose in commercial juices were measured at laboratory by standard AOAC method [9].

Preparation of Standard Mixture Solutions

Standard mixture solutions of three sugars available in mango juice, i.e., glucose, fructose and sucrose, were prepared. Eight different concentrations of glucose (0.5, 1.0, 1.5, 2.0, 2.5, 3, 5, 10 percent), fructose (0.5, 1.0, 1.5, 2.0, 2.5, 3, 5, 10 percent) and sucrose (7.0, 7.5, 8.0, 8.5, 9.0, 9.5, 10.0, 15.0 percent) were used to prepare artificial mixture solutions. These concentrations were considered here because the concentration of sugars in commercial juices varies among these ranges. Here "Orthogonal Experimental Design" was used to statistically maximize the information in the outputs. Thus, in total 64 mixtures were prepared with different concentrations of glucose, fructose and sucrose. According to the combination of concentrations from experimental design, the sugars were dissolved into de-ionized water to make solutions. Both mixture solutions and commercial juices were used in FTIR to get spectral data.

FTIR Measurements

Fourier Transformed Infrared (FTIR) spectrometer (Shimadzu, Model: IRAffinity1) connected to software of IR Solution Operating system (Version 1.40) was used to obtain FTIR spectra of samples. The samples were placed in contact with Attenuated Total Reflectance (ATR) element at controlled ambient temperature. FTIR spectra were collected in frequency 4000-650 cm^{-1} by co-adding 30 scans and at resolution of 4 cm^{-1} . Before every scan, a new reference air background spectrum was taken. These spectra were recorded as absorbance values at each data point in triplicate. The ATR plate was carefully cleaned in situ by wiping it with acetone, and dried with soft tissue before filling in with next sample.

Preprocessing of Spectral Data

The spectral data acquired from instrument contain spectra background information and noises which are interfered desired relevant quality attributes information. Interfering spectral parameters, such as light scattering, path length variations and random noise resulted from variable physical sample properties or instrumental effects need to be eliminated or reduced in order to obtain reliable, accurate and stable calibration models. Thus, it is very necessary to pre-process spectral data prior to modeling. Pre-processing methods can include averaging over spectra, which is used to reduce the number of wavelengths or to smooth the spectrum [9]. FTIR spectral data were preprocessed by de-noising them, and their efficiencies were assessed and compared here. For this three de-noising techniques, (1) Standard Normal Variate (SNV), (2) Multiplicative Scatter Correction (MSC) [10], and (3) Savitzky-Golay (S-G) filtering were used. Two de-noise efficiency measures, such as, Signal to Noise Ratio (SNR) and Root Mean Squared Error (RMSE) were used here.

All parts of a spectrum are not equally important to predict the considered variable. So it is necessary to find out which portion of the spectrum is contributing more to predict efficiently. Four different spectral range and full range were considered for analysis as spectral peaks and variation of absorbance could be observed in these regions of the spectra. The selected ranges are 3700-648 (full range), 3700-2880, 1800-1500 and 1500-952 wave number (cm^{-1}), where variations are noticed after plotting the spectral data against wave numbers.

Calibration Methods

Artificial Neural Network (ANN): An ANN is a data processing system based on the structure of the biological neural simulation by learning from the data generated experimentally or using validated models [11]. Application of ANN is a technique for data and knowledge processing, characterized by its analogy with a biological neuron [12], and can deal with nonlinear relationships between variables [13] and [14]. Use of ANN in food authentication studies has been gaining popularity among food scientists, food processing industries and food quality regulating authorities [15, 16].

A network consists of a sequence of layers with connections between successive layers. Data to the network is presented at input layer and the response of the network to the given data is produced in the output layer. There may be several layers between these two principal layers, which are called hidden layers. Finally, the use of a neural network approach to build a predictive model for a complex system does not require a statistician and domain expert to screen through every possible combination of variables. Thus, the neural network approach can dramatically reduce the time required to build a model.

In practical, Levenberg-Marquardt back-propagation neural network was used in the study. Number of hidden layers was 10 and the sigmoid activation function was used in each training. Training automatically stopped when generalization stops improving, as indicated by an increase in the root mean square error (RMSE) of the validation samples.

Principal Component Regression (PCR) and Partial Least Square Regression (PLSR): Absorbance of light by certain component against wave number is the main consideration for FTIR spectral data. In each spectrum there are huge number of absorbance values for each wave point. Each wave number or data point is considered as spectroscopic variable. These variables are huge in number and are mutually correlated. In this situation we cannot use Ordinary Least Square (OLS) regression to singularity problem. So, PCR and PLSR are playing a key role in spectroscopic data analysis. These two techniques extract a set of orthogonal factors called 'latent variables' from the predictors which have best predictive power. Here first 10 and 20 principal components were used respectively for prediction.

Calibration and Validation Datasets

Known concentrations of simple sugars and spectral data of 52 artificial mixture solutions were used to develop or train the models, 12 solutions for validate them. Finally, those of 15 commercial mango juices were used to test the efficiencies of developed models for predicting concentrations of glucose, fructose and sucrose in real samples of mango juice.

All the computer programs used to generate the results were written in MATLAB (version 8.1.0.604, The Math Works Inc., Natick, USA).

Results and Discussion

Spectral plot of mixture solutions of sugars and real mango juices are presented in Figure 1. Each spectra contains instrumental noise which have been removed first from the spectral data.

De-Noising of Spectral Data

De-noising performance of SNV, MSC and S-G filtering were assessed in different wave ranges, in different sets of data and in different calibration models, and the results are presented in Tables 1 to 6.

Prediction of Sugars Percentage by ANN with Raw and De-Noised Data

Prediction of glucose percentage by ANN with de-noised data: Prediction of glucose by Artificial Neural Network was

used with raw data and de-noised data with three methods, SNV, S-G filtering and MSC. It is evident that SNV treatment can not improve the prediction performance noticeably. Moreover, in some cases it is worse than that of raw data. Prediction is better with S-G filtering and MSC than that of raw data. Among these three de-noising techniques and raw data, S-G filtering shows the best results in terms of both RMSEP and R^2 in all spectral ranges and with all sets of data (training, validation and test) (Table 1).

Full spectral data and three different ranges of spectra have separately been calibrated with ANN for prediction of glucose. Calculated RMSEP and R^2 values show that prediction is best in the spectral range $1500-952\text{ cm}^{-1}$ among the selected ranges, and it is followed by full range, $3700-628\text{ cm}^{-1}$. Spectral range $3700-2880\text{ cm}^{-1}$ shows the next informative range and least one is $1800-1500\text{ cm}^{-1}$. That means, FTIR range $1500-952\text{ cm}^{-1}$ is the most informative region for prediction of glucose in mango juice by ANN. It is also noticed that among the de-noising techniques S-G filtering is the best in training, validation and test data sets (Table 1).

Prediction of fructose percentage by ANN with de-noised data:

Prediction performance of fructose by ANN with raw and de-noised data in selected spectral ranges are presented in Table 2. It is clear from the table that prediction is better with S-G filtering and MSC techniques of spectral data de-noising than raw data, but with SNV treatment, the performance is worse. S-G filtering shows the best results in terms of both RMSEP and R^2 . In case of fructose prediction, FTIR range $1500-952\text{ cm}^{-1}$ shows consistently maximum R^2 values for training, validation and test data.

Prediction of Sucrose Percentage by ANN with De-Noised data:

It is evident from Table 3 that prediction performance of sucrose by ANN is better after de-noised spectral data by S-G filtering and MSC than that of raw data, but with SNV treatment, the performance becomes worse. Among these three de-noised data and raw data S-G filtering shows the best results. Hence, in order to predict sucrose concentration in mango juice, ANN is an appropriate method but to get better result spectral data should be de-noised with S-G filtering.

Like glucose and fructose, $1500-952\text{ cm}^{-1}$ is the most informative among all selected spectral ranges for prediction of sucrose (Table 3).

To sum up, for prediction of glucose, fructose and sucrose by

Figure 1. FTIR Spectra of Mixture Solutions and Real Mango Juices.

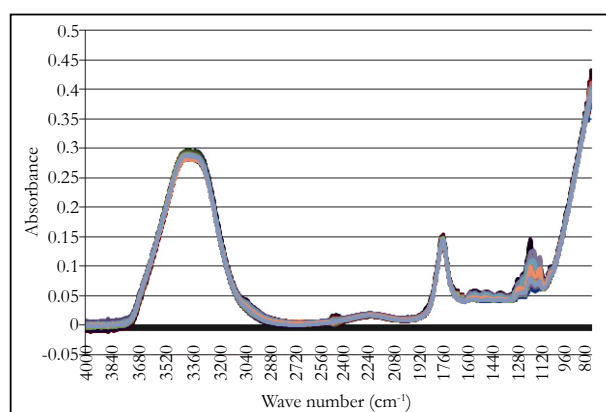


Table 1. Prediction of Glucose in Sugar Solutions and Real Mango Juice by ANN.

Spectral Range (cm ⁻¹)	Data sets (Sample number)	Raw data		SNV		S-G filtering		MSC	
		RMSEP	R ²	RMSEP	R ²	RMSEP	R ²	RMSEP	R ²
3700-628	Training (52)	1.3353	0.8152	3.88e ⁻¹⁰	1.0000	2.246e ⁻⁵	0.9995	0.1755	0.9962
	Validation (12)	0.9341	0.7619	0.8153	0.6003	0.0387	0.9982	0.5171	0.9084
	Test (15)	2.3400	0.5458	0.4208	0.8881	0.0509	0.9958	0.4564	0.9440
3700-2880	Training (52)	0.7151	0.9388	0.1072	0.9892	1.311e ⁻⁶	0.9982	9.561e ⁻⁶	0.9998
	Validation (12)	1.4038	0.8517	0.6133	0.6198	0.0404	0.9986	1.1368	0.7898
	Test (15)	1.2056	0.4808	0.8809	0.6159	0.0762	0.9872	1.7011	0.6149
1800-1500	Training (52)	1.0547	0.8473	0.6301	0.7298	0.0100	0.9898	7.759e ⁻⁴	0.9665
	Validation (12)	2.7169	0.2260	1.0061	0.3546	0.0519	0.9776	8.412e ⁻⁴	0.9657
	Test (15)	0.2688	0.2496	1.0331	0.4796	0.0566	0.9518	9.799e ⁻⁴	0.9454
1500-952	Training (52)	0.4079	0.9799	0.1273	0.9864	4.842e ⁻⁶	0.9998	0.1682	0.9964
	Validation (12)	1.0776	0.8199	0.3389	0.8921	0.0016	0.9996	0.6249	0.9324
	Test (15)	1.1528	0.9392	0.5447	0.7029	0.0031	0.9996	0.5941	0.9475

Table 2. Prediction of Fructose in Mixture Solutions and Real Mango Juice by ANN.

Spectral Range (cm ⁻¹)	Data sets (Sample number)	Raw data		SNV		S-G filtering		MSC	
		RMSEP	R ²	RMSEP	R ²	RMSEP	R ²	RMSEP	R ²
3700-648	Training (52)	1.6653	0.6438	0.2276	0.9580	6.008e ⁻⁵	0.9918	0.2175	0.9936
	Validation (12)	1.2626	0.7894	0.5493	0.4186	0.1183	0.9978	0.5029	0.9817
	Test (15)	1.4906	0.8272	0.8072	0.7066	0.0510	0.9868	0.4830	0.8312
3700-2880	Training (52)	0.3226	0.9843	0.5420	0.8425	0.0029	0.9990	0.2814	0.8849
	Validation (12)	1.4259	0.7032	0.7666	0.1569	0.0721	0.9874	1.0895	0.9004
	Test (15)	1.9061	0.6206	1.4727	0.7677	0.0548	0.9889	1.3363	0.5028
1800-1500	Training (52)	0.4281	0.9763	0.5417	0.8425	1.015e ⁻⁷	0.9928	0.8318	0.8481
	Validation (12)	2.3646	0.1696	0.7665	0.1569	0.0616	0.9810	1.0520	0.8255
	Test (15)	2.4287	0.0778	0.4655	0.7677	0.0819	0.9898	3.0095	0.0897
1500-952	Training (52)	0.3176	0.9837	0.7422	0.9399	1.243e ⁻⁴	0.9998	0.1942	0.9966
	Validation (12)	0.8508	0.9158	0.7790	0.2760	0.0027	0.9978	0.7371	0.8290
	Test (15)	2.9489	0.3874	1.4301	0.0967	0.0027	0.9952	0.6383	0.9537

Table 3. Prediction of Sucrose in Mixture Solutions and Real Mango Juice by ANN.

Spectral Range (cm ⁻¹)	Data sets (Sample number)	Raw data		SNV		S-G filtering		MSC	
		RMSEP	R ²	RMSEP	R ²	RMSEP	R ²	RMSEP	R ²
3700-648	Training (52)	0.0539	0.9994	0.2114	0.9759	2.063e ⁻⁴	0.9925	9.47e ⁻¹¹	0.9998
	Validation (12)	1.5695	0.6477	0.7355	0.3639	0.0592	0.9874	0.0933	0.9930
	Test (15)	2.1108	0.5006	0.4469	0.8547	0.0548	0.9910	0.1145	0.9823
3700-2880	Training (52)	0.3555	0.9712	0.0592	0.9970	4.782e ⁻⁵	0.9910	0.0374	0.9988
	Validation (12)	2.2688	0.2945	0.3629	0.1564	0.0583	0.9835	0.2672	0.8719
	Test (15)	2.7843	3.61e ⁻⁶	0.3362	0.4559	0.0361	0.9930	0.3859	0.8332
1800-1500	Training (52)	8.469e ⁻⁶	0.9998	0.2828	0.9485	3.637e ⁻⁷	0.9828	8.29e ⁻⁶	0.9998
	Validation (12)	3.3737	0.2705	0.8757	0.5857	0.0445	0.9392	0.4529	0.8001
	Test (15)	1.5588	0.2943	0.8958	0.1628	0.0938	0.9686	0.4024	0.7928
1500-952	Training (52)	0.8506	0.8815	0.2609	0.9708	1.876e ⁻⁴	0.9998	0.1153	0.9930
	Validation (12)	0.4664	0.9616	0.2848	0.4773	0.0479	0.9958	0.2429	0.9085
	Test (15)	0.8035	0.8925	0.5152	0.9247	0.0332	0.9966	0.2373	0.9065

ANN, FTIR wave number range 1500-952 cm⁻¹ shows constantly maximum R² values and minimum RMSEP values. Further, S-G filtering shows the best de-noising performance in all spectral ranges and with all sets of data (R²≈0.99).

Prediction of Sugars by PCR and PLSR with De-Noised Data

Principal Component Regression (PCR) and Partial Least Squares Regression (PLSR) are calibrated for quantification of glucose, fructose and sucrose with spectral values of artificial mixture solutions, and are validated with those of real mango juices in the study. Here calibrations are done with 10 and 20 latent variables, Principal Components (PCs). A comparative picture of predictive performance of these two statistical methods i.e., PCR and PLSR in terms of R² and RMSEP are shown in the Tables 4-6.

Prediction of simple sugars- glucose, fructose and sucrose is done after de-noising spectral data with SNV, S-G filtering and MSC. The PCR and PLSR models have been calibrated with full spectra and selected segments of spectra to select the best one among the alternatives.

It is noticed that, PLSR shows better results than PCR in all de-noising techniques and in all spectral range (Table 4). Best prediction performance is noticed in PLSR with 20 latent variables. S-G filtering performs better than other two de-noising techniques (SNV and MSC) and raw data. Full spectra (3700-648 cm⁻¹) and partial range 1500-952 cm⁻¹ show better prediction performance (R²≈0.99) than other spectral ranges. Due to management and

calculation simplicity, we are preferring shorter range, that is 1500-952 cm⁻¹ as better performing region.

For prediction of fructose concentration in mango juices, PLSR performs better than PCR, and the best results are obtained with 20 latent variables (Table 5). Among de-noising techniques, S-G filtering is the best on the basis of R² and RMSEP. All spectral ranges show very good prediction except 1800-1500 cm⁻¹ (R²≈0.98).

Prediction results of sucrose concentration show similar results as those of glucose and fructose by PCR and PLSR (Table 6). Here, de-noising technique, S-G shows the best performance, and spectral range 1500-952 cm⁻¹ is the most information region of the spectra (R²≈0.98).

To recapitulate, for prediction of glucose, fructose and sucrose, PLSR is better than PCR, and ANN performs better than PLSR. Savitzky-Golay (S-G) filtering shows the best performance to de-noise FTIR spectral data for preprocess than before calibration. Lastly, spectral range 1500-952 cm⁻¹ is the most informative part of the whole spectra for prediction of simple sugars in mango juice.

It is to be mentioned here that some amount of glucose, fructose and sucrose exists in ripe mango naturally. There is a base amount of these simple sugars in ripe mango of different varieties [1]. So amount of added sugars can easily be estimated by subtracting the base amount from the amount found in the proposed method.

Table 4. Prediction of Glucose Concentration by PCR and PLSR.

Spectral region (cm ⁻¹)	Pretreatment	PCR ₁₀		PLSR ₁₀		PCR ₂₀		PLSR ₂₀	
		R ²	RMSEP	R ²	RMSEP	R ²	RMSEP	R ²	RMSEP
3700-648	Raw data	0.8252	1.1191	0.9304	0.7060	0.8600	1.0014	0.9794	0.0668
	SNV	0.8325	0.4066	0.9445	0.2340	0.8763	0.3495	0.9896	0.0205
	S-G filtering	0.9904	0.0902	0.9886	0.0345	0.9993	0.0242	0.9998	0.0023
	MSC	0.9746	0.3613	0.9729	0.1905	0.9896	0.2310	0.9898	0.0292
3700-2880	Raw data	0.6322	1.6231	0.8819	0.3605	0.7181	1.4210	0.8867	0.6410
	SNV	0.6599	0.5795	0.9718	0.1340	0.7287	0.5175	0.9812	0.0232
	S-G filtering	0.9922	0.0512	0.9875	0.0293	0.9981	0.0405	0.9921	0.0117
	MSC	0.6055	1.4225	0.9676	0.2523	0.7098	1.1299	0.9701	0.0120
1800-1500	Raw data	0.3287	2.1930	0.7271	1.3980	0.4008	2.0718	0.8895	0.8897
	SNV	0.3556	0.7976	0.7530	0.4937	0.4458	0.7397	0.8661	0.3636
	S-G filtering	0.9879	0.1013	0.9561	0.0573	0.9973	0.0482	0.9828	0.0087
	MSC	0.5321	1.5491	0.7904	1.0369	0.5705	1.4841	0.9153	0.659
1500-952	Raw data	0.8476	1.0449	0.9607	0.5308	0.8865	0.9016	0.9797	0.0481
	SNV	0.8567	0.3761	0.9613	0.1954	0.3431	0.8807	0.9896	0.0190
	S-G filtering	0.9981	0.0396	0.9901	0.0278	0.9998	0.0228	0.9999	0.0022
	MSC	0.9253	0.6190	0.9819	0.2784	0.9479	0.5171	0.9837	0.0377

Table 5. PCR and PLSR for Prediction of Fructose Concentration.

Spectral region (cm ⁻¹)	Pretreatment	PCR ₁₀		PLSR ₁₀		PCR ₂₀		PLSR ₂₀	
		R ²	RMSEP	R ²	RMSEP	R ²	RMSEP	R ²	RMSEP
3700-648	Raw data	0.7906	1.2215	0.9043	0.8257	0.8019	1.1879	0.9793	0.0731
	SNV	0.7905	0.4548	0.9309	0.2612	0.8362	0.4022	0.9895	0.0228
	S-G filtering	0.9860	0.0719	0.9860	0.0386	0.9975	0.0304	0.9922	0.0027
	MSC	0.9739	0.3642	0.9729	0.1906	0.9895	0.2316	0.9898	0.0293
3700-2880	Raw data	0.5771	1.7354	0.8783	0.3933	0.7075	1.4432	0.8882	0.7310
	SNV	0.6064	0.6234	0.9771	0.1506	0.7378	0.5088	0.9875	0.0218
	S-G filtering	0.9803	0.0852	0.9670	0.0331	0.9966	0.0355	0.9813	0.0118
	MSC	0.6686	1.2987	0.9508	0.2161	0.7497	1.1286	0.9792	0.0771
1800-1500	Raw data	0.3870	0.7546	0.8946	1.3820	0.4981	1.8906	0.9158	0.7744
	SNV	0.3785	0.7833	0.7950	0.4498	0.5466	0.669	0.9501	0.2220
	S-G filtering	0.9868	0.0698	0.8957	0.0397	0.9958	0.0394	0.9273	0.0091
	MSC	0.4368	1.6930	0.7635	1.0971	0.4921	1.6076	0.8801	0.7811
1500-952	Raw data	0.8029	1.1848	0.9402	0.6529	0.8388	1.0713	0.9795	0.0618
	SNV	0.8078	0.4356	0.9399	0.2437	0.8030	0.4030	0.9895	0.0221
	S-G filtering	0.9941	0.0467	0.9980	0.0272	0.9977	0.0294	0.9992	0.0025
	MSC	0.9348	0.5762	0.9803	0.3164	0.9606	0.4476	0.9896	0.0449

Conclusion

It is evident from the results that FTIR spectral range 1500-982cm⁻¹ is the most informative region. Savitzky-Golay (S-G) filtering shows the best de-noising performance among the de-noising techniques considered in the study. Among the calibration models, Artificial Neural Network (ANN) demonstrates the highest prediction efficiency for glucose, fructose and sucrose.

Therefore, a novel analytical method is being proposed for quan-

tification of glucose, fructose and sucrose in commercial mango juice with the best alternatives of spectral data acquisition, pre-processing and calibration.

The proposed method is cost effective, time saving and do not generate any chemical waste as no standard chemical is used in this method. So, the method could save a huge amount of quality testing cost for mango juice producing companies, consumers and food quality regulating authorities.

Table 6. PCR and PLSR for Prediction of Sucrose Concentration.

Spectral region (cm ⁻¹)	Pretreatment	PCR ₁₀		PLSR ₁₀		PCR ₂₀		PLSR ₂₀	
		R ²	RMSEP	R ²	RMSEP	R ²	RMSEP	R ²	RMSEP
3700-648	Raw data	0.7488	1.0779	0.9240	0.5931	0.7959	0.9718	0.9791	0.0633
	SNV	0.7475	0.4993	0.9306	0.2617	0.8088	0.4344	0.9894	0.0242
	S-G filtering	0.9927	0.0495	0.9875	0.0289	0.9982	0.0245	0.9937	0.0018
	MSC	0.8800	0.0520	0.9525	0.0496	0.9812	0.0538	0.9800	0.0015
3700-2880	Raw data	0.4403	1.6091	0.9779	0.2367	0.5353	1.4662	0.9821	0.0408
	SNV	0.4162	0.7592	0.9852	0.1207	0.5717	0.6503	0.9952	0.0216
	S-G filtering	0.9922	0.0512	0.9975	0.0293	0.9967	0.0335	0.9989	0.0089
	MSC	0.9408	0.2244	0.9857	0.0608	0.9526	0.2009	0.9821	0.0131
1800-1500	Raw data	0.4282	1.6263	0.6810	1.2147	0.5066	1.5108	0.8636	0.7943
	SNV	0.3927	0.7743	0.6793	0.5627	0.4444	0.7407	0.8896	0.3301
	S-G filtering	0.9921	0.0518	0.9467	0.0334	0.9973	0.0301	0.9692	0.0061
	MSC	0.9614	0.1813	0.9026	0.1215	0.9678	0.1656	0.9130	0.0772
1500-952	Raw data	0.7199	1.1382	0.9407	0.5238	0.8299	0.8870	0.9779	0.0984
	SNV	0.7202	0.5256	0.9416	0.2397	0.8297	0.4100	0.9764	0.0510
	S-G filtering	0.9962	0.0356	0.9917	0.0281	0.9980	0.0257	0.9996	0.0025
	MSC	0.8882	0.3083	0.9724	0.1531	0.9320	0.2405	0.9896	0.0190

References

- [1]. Ara R, Motalab M, Uddin MN, Fakhruddin ANM, Saha BK (2014) Nutritional evaluation of different mango varieties available in Bangladesh. *Int Food Res J.* 21(6): 2169-2174.
- [2]. Tandel RK (2011) Sugar Substitutes: Health controversy over perceived benefits. *J Pharmacol Pharmacother.* 2(4): 236-243.
- [3]. Blanco-Gomis D, Fernandez-Rubio P, Gutierrez-Alvarez MD, Mangas-Alonso JJ (1998) Use of high-performance liquid chromatographic-chemometric techniques to differentiate apple juices clarified by microfiltration and ultrafiltration. *Analyst.* 123(1): 125-129.
- [4]. Reid LM, O'Donnell CP, Kelly JD, Downey G (2004) Preliminary Studies for the Differentiation of Apple Juice Samples by Chemometric Analysis of Solid-Phase Microextraction-Gas Chromatographic Data. *J Agricul Food Chem.* 52(23): 6891-6896.
- [5]. Helena R, Teresa S (2003) Determination of Saccharides in Multifloral Honey by means of HPLC. *J Apicultural Sci.* 47(2): 93-101.
- [6]. Bro R, Berg F, Thybo A, Andersen CM, Jorgensen BM, et al., (2002) Multivariate data analysis as a tool in advanced quality monitoring in the food production chain. *Trends in Food Sci Technol.* 13(6-7): 235-244.
- [7]. Iola FD, Barros A, Delgadillo I, Almeida C, Gil AM (2002) Application of FTIR Spectroscopy for the Quantification of Sugars in Mango Juice as a Function of Ripening. *J Agricul Food Chem.* 50(11): 3104-3111.
- [8]. Jha SN, Gunasekaran S (2010) Authentication of sweetness of mango juice using Fourier transform infrared-attenuated total reflection spectroscopy. *J Food Eng.* 101: 337-342.
- [9]. Horwitz W (2005) Official Methods of Analysis of AOC International. 18th (edn), Maryland, USA.
- [10]. Nicolai BM, Beullens K, Bobelyn E, Peirs A, Saeys W, et al., (2007) Non-destructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology Technol.* 46: 99-118.
- [11]. Rinnan A, Berg F, Engsen SB (2009) Review of the most common pre-processing techniques for near-infrared spectra. *TrAC.* 28(10): 1201-1222.
- [12]. Bhotmange M, Shastri P (2011) Application of Artificial Neural Networks to Food and Fermentation Technology, Artificial Neural Networks-Industrial and Control Engineering Applications. *Intech open sci.* 201-222.
- [13]. Otto M (1999) Chemometrics: Statistics and computer application in analytical chemistry. Hoboken, NJ: Wiley-VCH, 314.
- [14]. Lozano J, Santos JP, Aleixandre M, Sayago I, Gutierrez J, et al., (2006) Identification of typical wine aromas by means of electronic nose. *IEEE Sensors J.* 6(1): 173-178.
- [15]. Cozzolino D, Smyth HE, Lattey KA (2006) Combining mass spectrometry based electronic nose, visible-near infrared spectroscopy and chemometrics to assess the sensory properties of Australian Riesling wines. *Analytica Chimica Acta.* 563(1): 319-324.
- [16]. Goodacre R, Hammond D, Douglas BK (1997) Quantitative analysis of the adulteration of orange juice with using pyrolysis mass spectrometry and chemometrics. *J Analy Appl Pyrolysis.* 40: 135-158.
- [17]. Lopez-Diez CE, Bianchi G, Goodacre R (2003) Rapid Quantitative Assessment of the Adulteration of Virgin Olive Oils with Hazelnut Oils Using Raman Spectroscopy and Chemometrics. *J Agric Food Chem.* 51(21): 6145-6150.