Prediction of vinegar processing parameters with chemometric modelling of spectroscopic data

Cagri Cavdaroglu [1], Banu Ozen [2,*]

Izmir Institute of Technology, Department of Food Engineering, Urla-Izmir, Turkey

ARTICLE INFO

ABSTRACT

Spectroscopic methods have the advantages of being rapid and environmentally friendly and can be used in measurement and control of processing parameters during food production. It was aimed to predict several quality and chemical parameters of vinegar processing from UV-visible and mid-infrared spectroscopic profiles. Two processing lines of both traditional and submerged vinegar production from 2 separate grape varieties (green and red grapes) were monitored. Some of the important markers of the fermentation processes; pH, brix, total acidity, total flavonoid content, total and individual phenolic contents, organic acid, sugar, ethanol concentrations as well as UV-visible and mid-infrared spectra were obtained during both types of vinegar processing and quality and chemical parameters were predicted from spectroscopic data using chemometric methods. Individual UV-visible and mid-infrared spectral profiles along with low level of data fusion were used in building of chemometric prediction models. Accurate, reliable and robust prediction models (R^2cal and R^2val >0.9) were obtained for quality parameters mostly with combination of two spectroscopic datasets. Predictive models used for phenolic components were below average except for p-coumaric and syringic acids. Citric and acetic acids were the most accurately estimated ones among organic acids along with ethanol. Close agreements between reference and predicted values were obtained during the monitoring of changes of some quality parameters for vinegar fermentation process through rapid and simultaneous spectroscopic measurements.

1. Introduction

Vinegar production is a two-stage process: alcoholic and acetic acid fermentations. Sugar source is converted into ethanol and CO_2 in the first stage and fermentation takes place with the activity of *Saccharomyces cerevisiae* strains in anaerobic conditions. During the second stage of processing, acetic acid and water are produced from ethanol by acetic acid bacteria in aerobic conditions.

Vinegar is commonly produced with traditional and submerged fermentation techniques. Traditional vinegar processing involves fermentation by the microbial culture which forms a film on the surface. A relatively longer time of around 6–14 weeks is required for acetification of the must using this method [1]. Submerged fermentation, on the other hand, is a faster production technique. Fermentation takes place with the activity of acetic acid bacteria which is homogenously distributed in must [1]. Bioreactor is aerated from the bottom so that fermentation occurs not only on the surface but also throughout all

fermentation media. Therefore, this type of production allows fast conversion to acetic acid and high yield and is preferred as a commercial processing technique. Acetic acid at 8–9% levels can be obtained within 24–48 h after ethanol fermentation.

Vinegar composition mainly depends on raw material and production technique. Acetic acid and water constitute most of the vinegar; however, small amounts of organic acids, alcohol, phenolic compounds and amino acids are also present. Minor compounds are especially important for sensorial characteristics of this product.

Various spectroscopic methods have been used especially in the characterization and authentication of different types of vinegars [26–29] and use of these techniques for vinegar was reviewed in literature [5,25]. Spectroscopic methods have been also applied to predict the chemical compositional parameters of different types of food products. There are studies in literature which monitored the critical compositional parameters such as total acidity, sugar, acetic acid and ethanol contents at different stages of vinegar production and minor

---

components such as volatiles, phenolic profile and total phenol content were also determined throughout fermentation processes [2,7,15,37]. Spectral data evaluated with chemometric techniques allow the simultaneous estimation of the concentrations of chemical constituents of different types of fermented food products [4,10,9,11,19]. Therefore, in some studies, various spectroscopic profiles during vinegar production or only of final product were also collected to predict the quality and chemical parameters. Fourteen parameters including total acidity, volatile and non-volatile acids, organic acids, L-proline, dry matter, ash and chlorine contents of wine vinegar were successfully predicted from partial least square (PLS) regression models of near infrared (NIR) spectroscopic data [30]. Acetification process of vinegar produced from onion waste was followed with ethanol, acetic acid, biomass and NIR spectral measurements and these parameters were determined from the spectral data with PLS regression modelling [12]. NIR spectroscopy was also used in estimating the ethanol and acetic acid concentrations in culture broth samples obtained from rice vinegar fermentation [41]. In another study, Raman spectroscopy was used in monitoring grape vinegar production and, changes in glucose, fructose, ethanol and acetic acid concentrations were predicted with high coefficient of determination values through the evaluation of spectral data with PLS regression [39]. In this study, traditional and submerged fermentation types of vinegar production from two grape varieties were monitored with the determination of 22 quality and chemical parameters (brix, total phenolic content, total flavonoid content, titratable acidity, pH, and concentrations of citric acid, lactic acid, malic acid, succinic acid, tartaric acid, acetic acid, caffeic acid, catechin, epicatechin, coumaric acid, gallic acid, syringic acid, vanillic acid, ethanol, sucrose, glucose, fructose) along with the collection of UV-visible (UV-Vis) and Fourier transform infrared (FTIR) spectra. It was aimed to predict these quality and chemical parameters from spectral data using various chemometric techniques in order to determine several parameters simultaneously and rapidly during vinegar production.

2. Materials and methods

2.1. Vinegar production and sampling

Dried Sultaniye (white grape) and Alicante Bouchet (red grape) types of grapes were used separately in the production of vinegars. Samples from submerged culture fermentation were obtained from a commercial vinegar production line for these two grape varieties separately. Sampling was done at various times of alcoholic and acetic acid fermentations twice. 29 and 71 samples were collected during alcoholic and acetic acid fermentations, respectively.

Traditional type (surface fermentation) of vinegar processing was done with the same type of grapes separately and 2 batches were prepared for each grape variety. Grape musts obtained from a commercial vinegar processing plant were used as raw materials for this type of production. Mother of vinegar obtained during pre-trials were added to grape musts (18 Brix) and musts, in glass bottles covered with cotton cloths, were kept in a dark place. Sampling was done at 0th, 2nd, 4th, 6th, 10th, 15th, 20th, 25th, 30th and 40th days and a total of 40 samples were obtained during traditional production. In addition, 26 commercial vinegars were obtained from markets to widen the range of measured variables and to increase the number of the samples which are critical in building prediction models.

2.2. Chemical analysis

2.2.1. Brix, pH and total acidity measurements
pH of the samples was measured with a pH meter (WTW, Germany). Brix was determined with a digital refractometer (Isolab, Germany). Total acidity was measured with titration using NaOH [14] and expressed as volumetric percentage (% v/v).

2.2.2. Total phenolic and flavonoid contents
Total phenolic content (TPC) of the samples were determined with a spectrophotometric Folin-Ciocalteu assay adapted to microscale [24]. Results were reported as mg gallic acid/L. Total flavonoid content (TFC) was measured at 510 nm with a spectrophotometer [43] and expressed as mg catechin/L.

2.2.3. Phenolic profiles
Concentrations of individual phenolic compounds were determined according to a method described in the literature [38]. Samples were filtered through a syringe filter (0.45 μm, cellulose acetate) before chromatographic analysis. Then, they are injected into an HPLC-DAD system (Perkin Elmer 200, Waltham, MA, USA) according to the conditions given in the same reference. C18 column (250×4.6 mm, 5 μm, ACE, Aberdeen, Scotland) was used in the analyses. Phenolic contents were calculated from at least 5 points standard curves of catechin, epicatechin, gallic acid, caffeic acid, syringic acid, p-coumaric acid and vanillic acid. All phenolic standards were purchased from Sigma-Aldrich (Germany).

2.2.4. Organic acid, sugar and ethanol contents
Organic acid, sugar and ethanol concentrations of vinegars were determined simultaneously with an HPLC having refractive index detector (Agilent 1200, Santa Clara, CA, USA) according to a method in literature [6]. Aminex 87H column (300×7.8 mm, 9 μm, Bio-Rad Laboratories, Hercules, CA, USA) was used for the analysis. Acetic, citric, malic, tartaric and succinic acids, glucose, fructose, sucrose and ethanol concentrations were determined from standard curves. All standards were obtained from Sigma-Aldrich (Germany).

2.3. Spectroscopic analyses

2.3.1. UV-visible spectroscopy
UV-visible (UV-Vis) spectra of the samples were collected in 200–550 nm range with a Thermo Multiscan UV-Vis spectrophotometer (Thermofisher Scientific, Waltham, MA, USA).

2.3.2. Fourier transform infrared spectroscopy
Mid-infrared spectra of the samples were obtained with an FTIR spectrophotometer having a horizontal ZnSe-ATR accessory and a DTGS detector (Spectrum 100, Perkin Elmer, Waltham, MA, ABD) in 4000–800 cm^{-1} range. Measurements were performed with 128 scans and 4 cm^{-1} resolution. Spectra of air were taken as background before each sample reading.

2.4. Chemometric modelling

All chemometric analyses were performed with 'ropls package' (Version 3.12) in R [33]. 2/3 of the data were used for calibration and 1/3 was separated to validate the models, and 107 and 62 samples were used in the development of calibration and validation models, respectively. Stratified random sampling was applied prior to multivariate statistical analyses to choose calibration and validation samples [31]. For measured properties, every sample was split into subgroups based on percentiles and random sampling was done within these subgroups.

Chemometric models were constructed to predict the chemical parameters of the samples that were obtained during two types of vinegar production along with commercial vinegars from individual UV-Vis and FTIR spectra. FTIR and UV-Vis measurements contain absorbance values between 4000 and 800 cm^{-1} wavenumbers and 200–550 nm wavelengths of the samples, respectively. Low level data fusion with the combinations of two spectroscopic data were also used in model building. Complementary integration of homogeneous FTIR data with UV-Vis data was applied to increase descriptive power and to reduce information gaps [32].

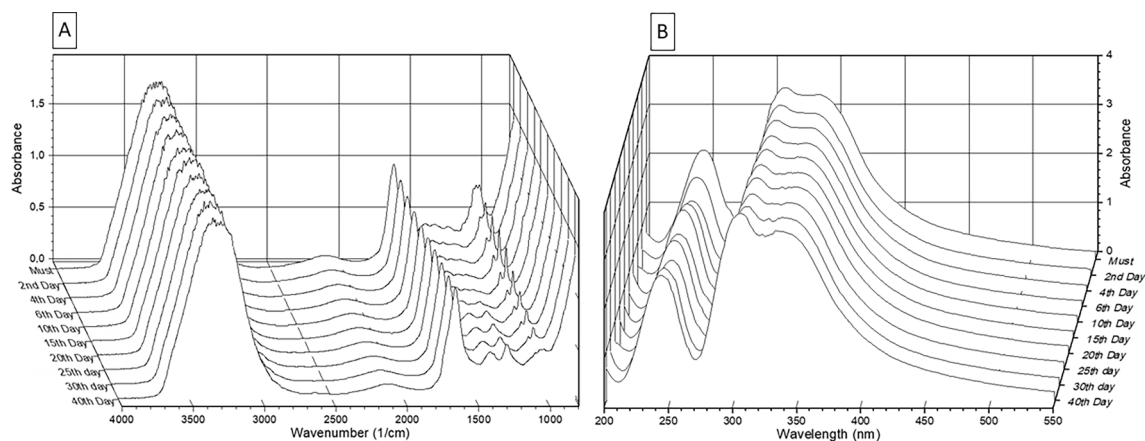Partial least square (PLS) and orthogonal partial least square (OPLS)

Fig. 1. FTIR (A) and UV-Vis (B) spectra of the samples collected during traditional vinegar processing of Sultaniye grape must.

regression methods were used to generate the prediction models. All spectroscopic data including two individual spectroscopic profiles along with their combinations were transformed with square, first, second and third derivative transformations, Savitzky-Golay filtering (SGF), standard normal variate (SNV) and multiplicative signal correction (MSC) methods before construction of prediction models for each variable. More information regarding the pre-processing techniques can be found in literature [20,23]. Fourteen models were generated for each parameter and the performance of these models were tested with the number of latent variables (LV), coefficient of determination for calibration (R^2cal), coefficient of determination for validation (R^2val), root mean square of error for calibration (RMSEC), root mean square of error for prediction (RMSEP) and residual predictive deviation (RPD) [36]. R^2 values close to 1 and small RMSE values relative to measurement ranges show the reliability of the models. RPD can be used as an indicator for the evaluation of a model's predictive ability. RPD value which is less than 1.5 indicates that the model's predictive capability is poor. Model is classified as average when the RPD value is between 1.5 and 2.0. RPD values between 2.0 and 2.5 shows that the model effect is relatively good and it is suitable for quantitative analysis. RPD values between 2.5 and 3.0 shows that the model is very effective and higher values than 3 indicates that the model has a very good prediction ability [42].

3. Results and discussion

Quality parameters and concentrations of several important components during grape vinegar production with two different techniques

(traditional and submerged culture fermentation) were determined using reference methods. Two different grape types and two production techniques along with commercial vinegar samples provided a wide range of parameters. Range and spread of measured values and number of analyzed samples are critical to obtain good prediction models. Reference results were compared with predicted results obtained from PLS and OPLS models developed by using FTIR and UV-Vis spectroscopic data along with their combinations. Several transformations were applied to all data before model building for each parameter as explained in Section 2.4 and only the results of the best models are presented here.

Sample UV-Vis and FTIR spectra obtained during vinegar production are shown in Fig. 1. As expected, both spectra have variations in the absorbance values of the peaks with respect to process stage due to reactions taking place throughout the processing. The peaks in 280–500 nm region of UV-Vis spectra are associated with phenolic compounds and organic acids [35,40]. As far as the FTIR spectra is concerned, major differences were observed in 1500–900 cm^{-1} region although all peaks varied somewhat with processing stage. Peak in 3800–2790 cm^{-1} region is attributed to –OH group of water and C—H stretching of acetic acid. In addition, 1300–1000 cm^{-1} is related with absorption due to organic acids while the peak in 1100–1000 cm^{-1} belongs to C-O stretching. Then, peaks at 1065–1030 cm^{-1} are associated with O—H and –CH$_2$ groups of sugars. Absorptions due to C=O stretching of aldehydes, -C-O and –OH groups of phenolic compounds take place in 1700–1600 cm^{-1} and 1800–900 cm^{-1} regions of FTIR spectra, respectively [10,26].

Table 1
Statistical parameters of predictive models for quality measurements.

Component	Dataset	Transformation*	Method	LV	R^2-Cal	R^2-Val	RMSEC	RMSEP	RPD
Brix	Combined	SNV	OPLS	9	0.98	0.98	0.47	0.48	2.65
	FTIR	SNV	PLS	7	0.98	0.97	0.47	0.62	1.92
	UV-Vis	Raw Data	PLS	6	0.97	0.97	0.61	0.65	1.72
TPC	Combined	Raw Data	OPLS	9	0.99	0.97	53.07	107.54	2.70
	FTIR	MSC	PLS	6	0.97	0.96	96.64	119.21	2.39
	UV-Vis	MSC	PLS	5	0.97	0.96	96.24	113.78	2.30
TFC	Combined	SNV	PLS	11	0.99	0.97	18.03	26.11	2.99
	FTIR	MSC	PLS	6	0.96	0.89	31.79	52.64	1.38
	UV-Vis	SGF	PLS	6	0.94	0.93	37.90	41.38	1.99
pH	Combined	Square	PLS	12	1.00	1.00	0.04	0.10	2.29
	FTIR	Square	PLS	10	1.00	1.00	0.06	0.10	2.48
	UV-Vis	SGF	OPLS	19	1.00	0.99	0.11	0.33	0.92
Titratable Acidity	Combined	Raw Data	PLS	9	0.99	0.99	0.36	0.27	5.35
	FTIR	Raw Data	OPLS	5	0.97	0.99	0.50	0.30	5.12
	UV-Vis	SGF	OPLS	20	0.94	0.54	0.73	2.09	0.91

* SNV: standard normal variate, MSC: multiplicative signal correction, SGF: Savitzky-Golay Filtering.
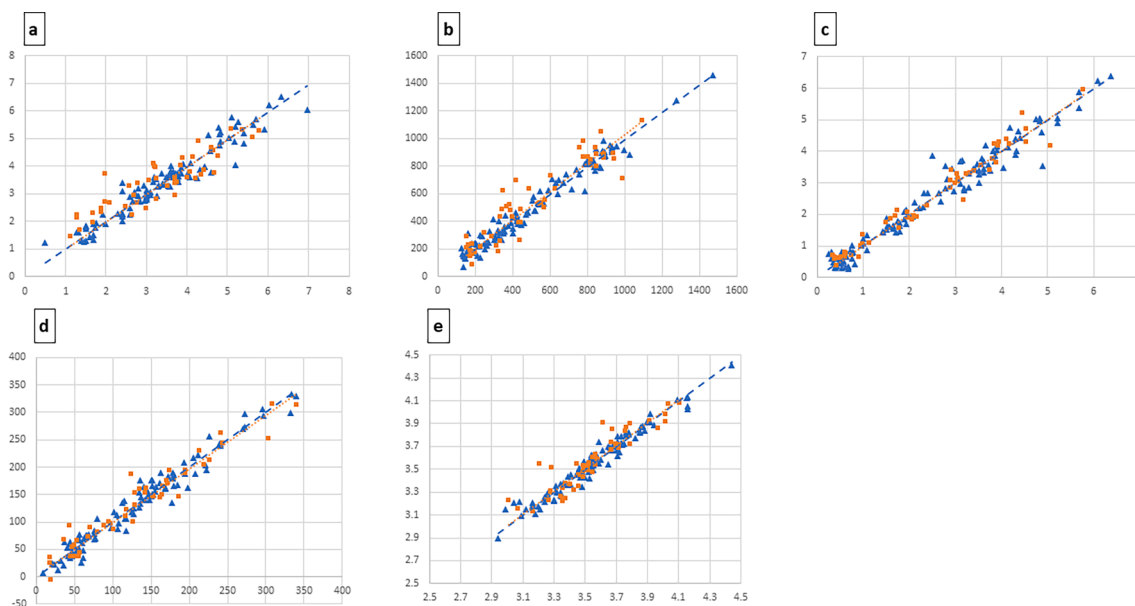
Fig. 2. Measured vs predicted plots constructed with the best chemometric regression models for a) brix, b) total phenol content (TPC), c) titratable acidity, d) total flavonoid content (TFC), and e) pH.

3.1. Prediction of quality parameters

Prediction results of five quality parameters (brix, TPC, TFC, pH and titratable acidity) are shown in Table 1. Brix values of the samples range between 0.5 and 31.2. PLS and OPLS regression analyses indicated that the use of only FTIR and UV-Vis data resulted in average models according to the performance criteria explained in Section 2.4; however, low level data fusion improved prediction vigor. Combination of FTIR and UV-Vis spectral data analyzed with OPLS after transformation with SNV generated the best results (R^2cal = 0.984, R^2val = 0.983, RMSEC = 0.468, RMSEP = 0.478, RPD = 2.652). TPC of the samples were determined as in the range of 120 – 3020 ppm. Similarly, combined dataset without any transformation provided the best predictive model (R^2cal = 0.992, R^2val = 0.969, RMSEC = 53.07, RMSEP = 107.5, RPD = 2.704), while the model predictions created with FTIR and UV-Vis datasets had average precisions. Although RMSEC values of TPC predictive models are a little bit high, measured TPC values are also high and comparison should be done considering measured values. A reason for high RMSEC values can be the relatively higher standard deviations of TPC measurements. However, these standard deviations are taken into account in RPD calculations and RPD and R^2 values of TPC model indicate very effective predictive ability of the data fusion model for this variable [42]. Maximum and minimum TFC values of the samples were 1.62 and 1500 ppm, respectively. TPC and TFC had wide ranges since both white and red grape types were used in the production. Although combined dataset estimated the closest values to the reference measurements (R^2cal = 0.986, R^2val = 0.973, RMSEC = 18.03, RMSEP = 26.11, RPD = 2.993) similar to the previous parameters, UV-Vis dataset seemed to be dominant over FTIR dataset in the combination model. Since UV-Vis spectroscopy is based on absorption of colored components its dataset provided good results for TPC and TFC predictions. Since grape juice has already acidic properties, pH values of the samples varied between 2.78 and 4.44. In the prediction of this parameter, UV-Vis dataset provided less contribution compared with the previous parameters. As a result, square transformed FTIR dataset with PLS regression had the most accurate results with R^2cal = 1, R^2val = 0.999, RMSEC = 0.055, RMSEP = 0.099, RPD = 2.481. Titratable acidity of vinegars results from the presence of different organic acids at different stages of fermentation. Late fermentation stages were dominated by acetic acid produced by the activity of acetic acid bacteria, while acidity is originated from fruit

itself at the beginning of the fermentation. Maximum and minimum titratable acidity values of the samples were recorded as 0.25 and 7.94, respectively. UV-Vis dataset was unsuccessful to create robust prediction model; however, FTIR dataset and FTIR dominated combined dataset analyzed with PLS and OPLS resulted in excellent predictive models (R^2cal = 0.986, R^2val = 0.991, RMSEC = 0.355, RMSEP = 0.273, RPD = 5.351). Low level data fusion was more successful in construction of predictive models for quality parameters except pH compared with individual spectroscopic data. FTIR spectroscopic data alone provided a better result for estimation of pH values of the samples. Although number of LV's are between 9 and 11 for these models, models were built using 3450 variables (3200 for FTIR and 250 for UV-Vis) with combination of two spectroscopic data sets. Graphs of measured vs. predicted values plotted using the best prediction models are shown in Fig. 2. As can be seen from these graphs and Table 1, very good agreements between measured and predicted values were obtained.

Several studies that used individual or combination of spectroscopic techniques with chemometric methods on vinegar samples are present in the literature. Some of these articles focused on commercial final products while the others aimed to monitor fermentation process. Soluble solids content and pH of white vinegars were determined using Vis/NIR data that was analyzed with least square support vector machine (LS-SVM) and PLS regression methods [3]. In another study, total acidity of traditional Chinese vinegars was predicted correctly using NIR data analyzed with non-linear regression technique [8]. Mid-IR spectroscopy connected with flow lines was used to determine acidity on a group of samples containing wine, cherry, apple and balsamic vinegars [21]. Analysis of data with parallel factor analysis (PARAFAC) and PLS regression provided excellent predictive models. TPC of apple, rice, grape, pomegranate, balsamic, white, rose and red wine vinegars was determined using FTIR spectroscopy data and PLS regression technique by [16]. As in the examples of these studies in literature, successful models also were obtained for the estimation of the quality parameters during vinegar processing in this study.

3.2. Prediction of phenolic compounds

In addition to TPC, estimation of phenolic compounds, that are known to be in vinegar, were studied. Reference values for those phenolic compounds were measured with HPLC. Prediction results of

Table 2
Predictive models for phenolic compounds content.

Component	Dataset	Transformation*	Method	LV	R^2-Cal	R^2-Val	RMSEC	RMSEP	RPD
Gallic Acid	Combined	Raw Data	OPLS	11	0.97	0.82	0.46	1.23	1.24
	FTIR	SNV	OPLS	6	0.89	0.81	0.90	1.27	1.07
	UV-Vis	Square	PLS	5	0.84	0.79	1.10	1.34	0.97
Catechin	Combined	SGF	PLS	11	0.92	0.61	0.53	1.07	0.80
	FTIR	SGF	PLS	7	0.84	0.64	0.74	1.02	0.65
	UV-Vis	SGF	OPLS	8	0.74	0.69	0.95	0.98	0.68
Epicatechin	Combined	SNV	PLS	16	0.99	0.78	0.61	2.50	1.32
	FTIR	Square	PLS	9	0.91	0.72	1.44	2.83	0.85
	UV-Vis	Square	PLS	7	0.83	0.68	2.00	3.01	1.02
p-coumaric Acid	Combined	SNV	PLS	14	0.98	0.87	0.10	0.32	1.84
	FTIR	Savitzky-Golay Filtering	OPLS	11	0.97	0.60	0.15	0.57	1.06
	UV-Vis	First Derivative	PLS	5	0.93	0.87	0.21	0.32	1.92
Caffeic Acid	Combined	Square	OPLS	7	0.85	0.77	1.83	2.17	1.36
	FTIR	SGF	OPLS	15	0.99	0.53	0.52	3.67	1.12
	UV-Vis	SGF	PLS	7	0.81	0.80	2.04	2.08	1.45
Vanillic Acid	Combined	MSC	OPLS	12	0.98	0.31	0.10	0.76	0.79
	FTIR	SGF	PLS	8	0.93	0.48	0.22	0.64	0.67
	UV-Vis	Third Derivative	PLS	5	0.96	0.37	0.16	0.72	0.67
Syringic Acid	Combined	SGF	PLS	10	0.97	0.86	0.06	0.15	1.81
	FTIR	Raw Data	PLS	6	0.91	0.83	0.11	0.16	1.38
	UV-Vis	SGF	OPLS	7	0.85	0.54	0.14	0.27	0.89

* SNV: standard normal variate, MSC: multiplicative signal correction, SGF: Savitzky-Golay Filtering.

Table 3
Predictive models for sugars, ethanol and organic acids.

Component	Dataset	Transformation*	Method	LV	R^2-Cal	R^2-Val	RMSEC	RMSEP	RPD
Citric Acid	Combined	SNV	OPLS	11	0.97	0.85	131.32	311.27	1.99
	FTIR	MSC	OPLS	4	0.77	0.79	343.83	363.42	1.54
	UV-Vis	SGF	PLS	7	0.81	0.80	313.45	353.52	1.46
Lactic Acid	Combined	Square	PLS	12	0.99	0.77	443.84	1947.51	1.05
	FTIR	Square	OPLS	13	0.99	0.70	362.79	2237.19	1.07
	UV-Vis	SGF	OPLS	16	0.93	0.49	987.23	3311.03	0.90
Malic Acid	Combined	SNV	OPLS	17	0.99	0.67	88.85	555.39	0.94
	FTIR	Square	PLS	11	0.97	0.51	155.75	694.85	0.82
	UV-Vis	SNV	PLS	6	0.86	0.64	353.05	564.47	0.77
Succinic Acid	Combined	MSC	PLS	9	0.91	0.79	524.75	874.39	1.05
	FTIR	Square	PLS	13	0.99	0.66	201.98	1110.15	0.89
	UV-Vis	SGF	OPLS	9	0.89	0.84	578.77	765.58	1.10
Tartaric Acid	Combined	SNV	PLS	14	0.99	0.87	149.50	667.07	1.46
	FTIR	Raw Data	PLS	8	0.97	0.74	285.33	1013.74	1.23
	UV-Vis	SGF	OPLS	10	0.93	0.64	461.66	1300.20	1.00
Acetic Acid	Combined	Raw Data	PLS	8	0.98	0.90	0.37	0.95	1.92
	FTIR	MSC	PLS	6	0.98	0.91	0.40	0.91	1.94
	UV-Vis	SGF	OPLS	15	0.90	0.53	0.88	2.33	0.96
Sucrose	Combined	Square	PLS	13	0.99	0.49	11.69	84.27	0.77
	FTIR	Square	PLS	12	0.99	0.45	12.43	96.17	0.86
	UV-Vis	Square	OPLS	6	0.87	0.31	39.95	97.76	0.76
Glucose	Combined	SNV	PLS	11	0.93	0.28	276.76	978.03	0.88
	FTIR	MSC	OPLS	8	0.92	0.36	301.81	879.08	0.79
	UV-Vis	Square	OPLS	12	0.86	0.10	396.36	956.77	0.74
Fructose	Combined	SNV	OPLS	10	0.92	0.26	610.52	2827.76	0.83
	FTIR	Raw Data	OPLS	10	0.99	0.16	248.66	3019.96	0.71
	UV-Vis	Raw Data	OPLS	13	0.92	0.34	604.57	2622.50	0.85
Ethanol	Combined	MSC	OPLS	9	0.89	0.84	0.31	0.45	1.84
	FTIR	SNV	PLS	5	0.88	0.82	0.40	0.39	2.08
	UV-Vis	First Derivative	PLS	6	0.92	0.37	0.26	1.47	1.06

* SNV: standard normal variate, MSC: multiplicative signal correction, SGF: Savitzky-Golay Filtering.

individual phenolic compounds are shown in Table 2. Concentration value ranges were measured as 0 – 30 ppm for gallic acid, 0–60 ppm for catechin, 0 – 15 ppm for epicatechin, 0 – 10 ppm for coumaric acid, 0 – 40 ppm for caffeic acid, 0 – 15 ppm for vanillic acid and 0 – 3.5 ppm for syringic acid. Similar phenolic compounds were determined in studies performed with grape vinegars in the literature [18,22] and the concentrations of these compounds are function of the grape type and processing type and processing stage. Although usage of combined
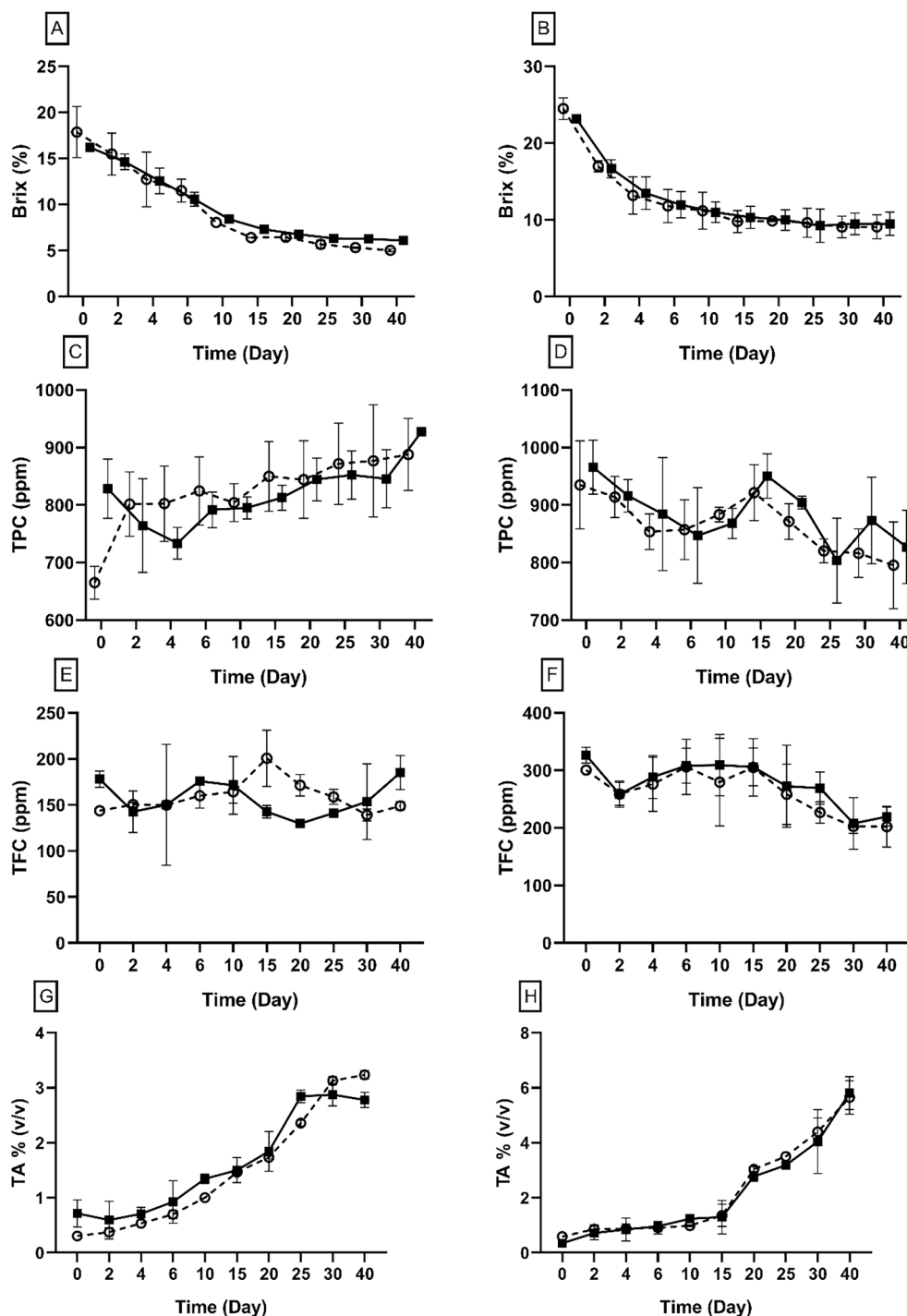
Fig. 3. Changes in quality parameters during surface fermentation process of vinegar with respect to time. A, C, E and G vinegar production with Sultaniye grapes and B, D, F and H are vinegar production with Alicante Bouchet grapes. Solid and dashed lines represent measured and predicted values, respectively.

dataset was more successful in the prediction of quality parameters, a generalization for prediction of phenolic compounds is not possible. Combined dataset with SGF resulted in the most successful prediction model for syringic acid, while first derivative transformation of UV-Vis dataset with PLS regression was preferable for p-coumaric acid. FTIR dataset alone did not produce any model better than UV-Vis data for the phenolic compounds. Combination models for gallic acid and epicatechin have high R^2cal (0.97 and 0.99) and average R^2val (0.82 and 0.78) values; however, their RPD values are not satisfactory. Expanded uncertainty [34] for gallic acid was calculated as 2.23 while RMSEP value for this parameter was 1.23. None of the datasets produced any

reliable model for catechin, caffeic acid and vanillic acid. Since these components are minor compounds of food matrix, statistical analyses resulted in average and below average success rates for predictive models.

3.3. Prediction of organic acids and sugars

Sugars and organic acids are the major components of fruits. Concentrations of these compounds and ethanol, produced during fermentation, are important parameters to monitor the process. Although total titratable acidity includes sum of individual organic acid concentration,

source of acidity may have importance for several processes such as wine fermentation. Obtained spectral data was also analyzed in order to test predictive capabilities in terms of sugars, acids and ethanol. Reference values for organic acids, sugars and ethanol were measured with HPLC. Prediction results of these compounds are shown in Table 3. Measured values ranged between 1 and 150 ppm for citric acid, 100–10,000 ppm for lactic acid, 4–2500 ppm for malic acid, 100–4500 ppm for succinic acid, 0–4730 ppm for tartaric acid, 0–7.14 % (v/v) for acetic acid, 0–300 ppm for sucrose, 0–3500 ppm for glucose, 0–10,000 ppm for fructose and 0–8 % (v/v) for ethanol. Although amounts of organic acids varied with grape type and processing, same type of organic acids were also determined in studies performed with grape vinegars [13,18,22]. Combined dataset produced better results for citric acid (R^2cal = 0.97, R^2val = 0.85) and tartaric acid (R^2cal = 0.99, R^2val = 0.87) prediction; nevertheless, FTIR data was favorable for estimation of ethanol (R^2cal = 0.88, R^2val = 0.82) and acetic acid (R^2cal = 0.98, R^2val = 0.91) concentrations. RMSEP value for citric acid was determined as 311, while expanded uncertainty was 568. UV-Vis dataset did not create any preferable predictive models. In the prediction of lactic acid, malic acid, succinic acid sucrose, fructose and glucose concentrations; none of the datasets were successful enough. Since these components can exist in very small amounts depending on the fermentation stage, concentrations in the data range are not well distributed and this causes generation of poor prediction models for some compounds. In literature, Vis/NIR data and various multivariate statistical analysis techniques were used in combination to predict organic acid content of plum vinegars and LS-SVM was determined as the most precise technique [17] and better prediction models using variable selection were developed for acetic, tartaric and lactic acids compared to current study.

Spectroscopic methods combined with chemometric techniques can provide opportunities to determine several quality parameters of food products simultaneously, rapidly and easily. In the current study, successful results were obtained for the estimation of brix, pH, titratable acidity, TPC and TFC along with average predictions of ethanol, acetic acid, citric acid, p-coumaric acid and syringic acid and, mostly combination of FTIR and UV-vis data provided better predictions.

Some of the most successful models (pH, titratable acidity, TPC and TFC) were used in predicting the changes during vinegar production. For this purpose, quality parameters of vinegar samples of both grape varieties which were collected during the production with surface fermentation technique are compared with the predicted values (Fig. 3). As can be seen from the figure, quite close agreements between predicted and measured values especially for pH and titratable acidity are observed. There are some deviations in TPC and TFC estimations. As can be seen from the plots (Fig. 3), particularly TPC measurements have relatively higher standard deviations. Therefore, deviations in predictions of these variables can be related with higher variability in measurements. Despite this, prediction models for quality variables can be considered as quite effective in monitoring the vinegar processing and can be used in monitoring of vinegar process.

4. Conclusion

In this study, various quality parameters, phenolic compounds, organic acid and sugar profiles of vinegars produced from Sultaniye and Alicante grape varieties by submerged and surface fermentation techniques are estimated from FTIR and UV-Vis spectral data in combination with PLS and OPLS regression analyses.

Spectral data and chemometric methods are successful in prediction of total amount of sugars, phenolics, flavonoids and organic acids. However, concentration of individual components which are portion of total sugar, phenolics or organic acids cannot be predicted with high precision. Successful results showed that FTIR and UV-Vis spectral data analyzed with chemometrics have potential to be cheap, non-hazardous, and fast methods in order to monitor vinegar fermentation processes. Simultaneous analyses of these parameters would provide better control

of quality during fermentation and also can be helpful in determining the authenticity of the product.

CRediT authorship contribution statement

Cagri Cavdaroglu: Conceptualization, Investigation, Methodology, Software, Formal analysis, Writing – original draft, Visualization. **Banu Ozen:** Conceptualization, Supervision, Resources, Project administration, Funding acquisition, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Adams, M. R. 1998. Vinegar. In B. J. B. Wood (Ed.), Microbiology of Fermented Foods (pp. 1–44). Springer US. https://doi.org/10.1007/978-1-4613-0309-1_1.

[2] W. Andlauer, C. Stumpf, P. Furst, Influence of the acetification process on phenolic compounds, Journal of Agricultural and Food Chemistry 48 (8) (2000) 3533–3536, https://doi.org/10.1021/jf000010j.

[3] Y. Bao, F. Liu, W. Kong, D.-W. Sun, Y. He, Z. Qiu, Measurement of Soluble Solid Contents and pH of White Vinegars Using VIS/NIR Spectroscopy and Least Squares Support Vector Machine, Food and Bioprocess Technology 7 (1) (2014) 54–61, https://doi.org/10.1007/s11947-013-1065-0.

[4] Canal, C., Ozen, B. 2017. Monitoring of Wine Process and Prediction of Its Parameters with Mid-Infrared Spectroscopy. Journal of Food Process Engineering, 40(1), e12280. https://doi.org/https://doi.org/10.1111/jfpe.12280.

[5] C. Cavdaroglu, B. Ozen, Authentication of Vinegars with Targeted and Non-targeted Methods, Food Reviews International 1–18 (2021), https://doi.org/10.1080/87559129.2021.1894169.

[6] M. Castellari, A. Versari, U. Spinabelli, S. Galassi, A. Amati, An Improved HPLC Method for the Analysis of Organic Acids, Carbohydrates, and Alcohols in Grape Musts and Wines, Journal of Liquid Chromatography & Related Technologies 23 (13) (2000) 2047–2056, https://doi.org/10.1081/JLC-100100472.

[7] C. Cejudo-Bastante, E. Durán-Guerrero, C. García-Barroso, R. Castro-Mejías, Volatile compounds, polyphenols and sensory quality in the production of tomato vinegar, Journal of Food Nutrition and Research 5 (2017) 391–398.

[8] Q. Chen, J. Ding, J. Cai, J. Zhao, Rapid measurement of total acid content (TAC) in vinegar using near infrared spectroscopy based on efficient variables selection algorithm and nonlinear regression tools, Food Chemistry 135 (2) (2012) 590–595, https://doi.org/10.1016/j.foodchem.2012.05.011.

[9] D. Cozzolino, R.G. Dambergs, L. Janik, W.U. Cynkar, M. Gishen, Analysis of Grapes and Wine by near Infrared Spectroscopy, Journal of Near Infrared Spectroscopy 14 (5) (2006) 279–289, https://doi.org/10.1255/jnirs.679.

[10] R. Croce, C. Malegori, P. Oliveri, I. Medici, A. Cavaglioni, C. Rossi, Prediction of quality parameters in straw wine by means of FT-IR spectroscopy combined with multivariate data processing, Food Chemistry 305 (2020) 125512, https://doi.org/10.1016/j.foodchem.2019.125512.

[11] B. Finn, L.M. Harvey, B. McNeil, Near-infrared spectroscopic monitoring of biomass, glucose, ethanol and protein content in a high cell density baker's yeast fed-batch bioprocess, Yeast 23 (7) (2006) 507–517, https://doi.org/10.1002/yea.1371.

[12] J.M. González-Sáiz, I. Esteban-Díez, C. Sánchez-Gallardo, C. Pizarro, Monitoring of substrate and product concentrations in acetic fermentation processes for onion vinegar production by NIR spectroscopy: value addition to worthless onions, Analytical and Bioanalytical Chemistry 391 (8) (2008) 2937–2947, https://doi.org/10.1007/s00216-008-2186-6.

[13] M. Hajimahmoodi, M. Khanavi, O. Sadeghpour, M.R.S. Ardekani, F.Z. Mazde, M.S. Khoddami, A.M. Ranjbar, Application of organic acid based artificial neural network modeling for assessment of commercial vinegar authenticity, Food Analytical Methods 9 (2016) 3451–3459, https://doi.org/10.1007/s12161-016-0510-x.

[14] International Organisation of Vine and Wine. 2000. OENO 52-2000 Wine vinegars – determination of total acidity content. Compendium of Methods of Analysis of Wine Vinegars. https://www.oiv.int/public/medias/2697/oeno-52-2000.pdf.

[15] Y. Jo, J.W. Han, D.-L. Min, Y.E. Lee, Y.-J. Choi, S. Lim, Optimization of acetic acid fermentation for producing vinegar from extract of jujube (Zizyphus jujuba Mill.) fruits, Korean Journal of Food Science and Technology 47 (6) (2015) 711–718.

[16] P. Kadiroğlu, FTIR spectroscopy for prediction of quality parameters and antimicrobial activity of commercial vinegars with chemometrics, Journal of The Science of Food and Agriculture 98 (11) (2018) 4121–4127, https://doi.org/10.1002/jsfa.2018.98.issue-1110.1002/jsfa.8929.

[17] F. Liu, Y. He, Application of successive projections algorithm for variable selection to determine organic acids of plum vinegar, Food Chemistry 115 (4) (2009) 1430–1436, https://doi.org/10.1016/j.foodchem.2009.01.073.

[18] Q. Liu, G.Y. Tang, C.N. Zhao, R.Y. Gan, H.B. Li, Antioxidant activities, phenolic profiles, and organic acid contents of fruit vinegars, Antioxidants 8 (2019) 78, https://doi.org/10.3390/antiox8040078.

[19] M. Loudiyi, H.T. Temiz, A. Sahar, M. Haseeb Ahmad, O. Boukria, A. Hassoun, A. Aït-Kaddour, Spectroscopic techniques for monitoring changes in the quality of milk and other dairy products during processing and storage, Critical Reviews in Food Science and Nutrition 1–25 (2020), https://doi.org/10.1080/10408398.2020.1862754.

[20] P. Mishra, A. Biancolillo, J.M. Roger, F. Marini, D.N. Rutledge, New data preprocessing trends based on ensemble of multiple preprocessing techniques, TrAC Trends in Analytical Chemistry 132 (2020) 116045, https://doi.org/10.1016/j.trac.2020.116045.

[21] J. Moros, F.A. Iñón, S. Garrigues, M. de la Guardia, Determination of vinegar acidity by attenuated total reflectance infrared measurements through the use of second-order absorbance-pH matrices and parallel factor analysis, Talanta 74 (4) (2008) 632–641, https://doi.org/10.1016/j.talanta.2007.06.046.

[22] R. Natera, R. Castro, M. de Valme García-Moreno, M.J. Hernández, C. García-Barroso, Chemometric studies of vinegars from different raw materials and processes of production, Journal of Agricultural and Food Chemistry 51 (2003) 3345–3351, https://doi.org/10.1021/jf021180u.

[23] P. Oliveri, C. Malegori, R. Simonetti, M. Casale, The impact of signal pre-processing on the final interpretation of analytical outcomes–A tutorial, Analytica Chimica Acta 1058 (2019) 9–17, https://doi.org/10.1016/j.aca.2018.10.055.

[24] I. Ozturk, O. Caliskan, F. Tornuk, N. Ozcan, H. Yalcin, M. Baslar, O. Sagdic, Antioxidant, antimicrobial, mineral, volatile, physicochemical and microbiological characteristics of traditional home-made Turkish vinegars, LWT-Food Scienceand Technology 63 (1) (2015) 144–151, https://doi.org/10.1016/j.lwt.2015.03.003.

[25] R. Ríos-Reina, J.M. Camiña, R.M. Callejón, S.M. Azcarate, Spectralprint techniques for wine and vinegar characterization, authentication and quality control: Advances and projections, TrAC Trends in Analytical Chemistry 134 (2021) 116121, https://doi.org/10.1016/j.trac.2020.116121.

[26] R. Ríos-Reina, R.M. Callejón, C. Oliver-Pozo, J.M. Amigo, D.L. García-González, ATR-FTIR as a potential tool for controlling high quality vinegar categories, Food Control 78 (2017) 230–237, https://doi.org/10.1016/j.foodcont.2017.02.065.

[27] R. Ríos-Reina, S.M. Azcarate, J. Camiña, R.M. Callejón, J.M. Amigo, Application of hierarchical classification models and reliability estimation by bootstrapping, for authentication and discrimination of wine vinegars by UV–vis spectroscopy, Chemometrics and Intelligent Laboratory Systems 191 (2019) 42–53, https://doi.org/10.1016/j.chemolab.2019.06.001.

[28] R. Ríos-Reina, S.M. Azcarate, J.M. Camiña, R.M. Callejón, Sensory and spectroscopic characterization of Argentinean wine and balsamic vinegars: A comparative study with European vinegars, Food Chemistry 323 (2020) 126791, https://doi.org/10.1016/j.foodchem.2020.126791.

[29] R. Ríos-Reina, S.M. Azcarate, J. Camiña, R.M. Callejón, Assessment of UV– visible spectroscopy as a useful tool for determining grape-must caramel in high-quality wine and balsamic vinegars, Food Chemistry 323 (2020) 126792, https://doi.org/10.1016/j.foodchem.2020.126792.

[30] M.J. Sáiz-Abajo, J.M. González-Sáiz, C. Pizarro, Prediction of organic acids and other quality parameters of wine vinegar by near-infrared spectroscopy, A feasibility study. *Food Chemistry* 99 (3) (2006) 615–621, https://doi.org/10.1016/j.foodchem.2005.08.006.

[31] Särndal C. E., Swensson, B., Wretman, J. 2003. In Model assisted survey sampling (pp.100–109). Springer.

[32] M. Schmitt, X.X. Zhu, Data fusion and remote sensing: An ever-growing relationship, IEEE Geoscience and Remote Sensing Magazine 4 (4) (2016) 6–23, https://doi.org/10.1109/MGRS.2016.2561021.

[33] E.A. Thévenot, A. Roux, Y. Xu, E. Ezan, C. Junot, Analysis of the Human Adult Urinary Metabolome Variations with Age, Body Mass Index, and Gender by Implementing a Comprehensive Workflow for Univariate and OPLS Statistical Analyses, Journal of Proteome Research 14 (8) (2015) 3322–3335, https://doi.org/10.1021/acs.jproteome.5b00354.

[34] T. Tomić, N.U. Nasipak, S. Babić, Estimating measurement uncertainty in high-performance liquid chromatography methods, Accreditation and Quality Assurance 17 (3) (2012) 291–300, https://doi.org/10.1007/s00769-011-0872-0.

[35] J.S. Torrecilla, R. Aroca-Santos, J.C. Cancilla, G. Matute, Linear and non-linear modeling to identify vinegars in blends through spectroscopic data, LWT – Food Science and Technology 65 (2016) 565–571, https://doi.org/10.1016/j.lwt.2015.08.027.

[36] A. Tugnolo, V. Giovenzana, C. Malegori, P. Oliveri, A. Casson, M. Curatitoli, R. Guidetti, R. Beghi, A reliable tool based on near-infrared spectroscopy for the monitoring of moisture content in roasted and ground coffee: A comparative study with thermogravimetric analysis, Food Control 130 (2021) 108312, https://doi.org/10.1016/j.foodcont.2021.108312.

[37] C. Ubeda, R.M. Callejón, C. Hidalgo, M.J. Torija, A.M. Troncoso, M.L. Morales, Employment of different processes for the production of strawberry vinegars: Effects on antioxidant activity, total phenols and monomeric anthocyanins, LWT – Food Science and Technology 52 (2) (2013) 139–145, https://doi.org/10.1016/j.lwt.2012.04.021.

[38] O. Uncu, B. Ozen, Geographical differentiation of a monovarietal olive oil using various chemical parameters and mid-infrared spectroscopy, Analytical Methods 8 (24) (2016) 4872–4880.

[39] R.S. Uysal, E.A. Soykut, I.H. Boyaci, A. Topcu, Monitoring multiple components in vinegar fermentation using Raman spectroscopy, Food Chemistry 141 (4) (2013) 4333–4343, https://doi.org/10.1016/j.foodchem.2013.06.122.

[40] O. Yalçın, C. Tekgündüz, M. Öztürk, E. Tekgündüz, Investigation of the traditional organic vinegars by UV–VIS spectroscopy and rheology techniques, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 246 (2021), 118987, https://doi.org/10.1016/j.saa.2020.118987.

[41] T. Yano, T. Aimi, Y. Nakano, M. Tamai, Prediction of the concentrations of ethanol and acetic acid in the culture broth of a rice vinegar fermentation using near-infrared spectroscopy, Journal of Fermentation and Bioengineering 84 (5) (1997) 461–465, https://doi.org/10.1016/S0922-338X(97)82008-9.

[42] C. Zhang, F. Liu, W. Kong, Y. He, Application of Visible and Near-Infrared Hyperspectral Imaging to Determine Soluble Protein Content in Oilseed Rape Leaves, In *Sensors* 15 (7) (2015) 16576–16588, https://doi.org/10.3390/s150716576.

[43] J. Zhishen, T. Mengcheng, W. Jianming, The determination of flavonoid contents in mulberry and their scavenging effects on superoxide radicals, Food Chemistry 64 (4) (1999) 555–559, https://doi.org/10.1016/S0308-8146(98)00102-2.