Classification of Plants Medicine Species from Algerian Regions using UV Spectroscopy, HPLC Chromatography, and Chemometrics Analysis

Soumeia Zeghoud[1,2,3], Bachir Ben Seghir[1,3,4], Imane Kouadri[3,4,5*], Hadia Hemmami[1,3], Ilham Ben Amor[1,3], Ali Tliba[2], Sadok Nani[2], Chinaza Godswill Awuchi[6], Mohammed Messaoudi[7,8] and Abdelkrim Rebiai[3,7]

[1]Department of Process Engineering and Petrochemical, Faculty of Technology, University of El Oued, El Oued 39000, Algeria

[2]Laboratory Valorization and Technology of Saharan Resources (VTRS), University of El Oued, El Oued 39000, Algeria

[3]Renewable Energy Development unit in Arid Zones (UDERZA), University of El Oued, El Oued 39000, Algeria

[4]Laboratory of Industrial Analysis and Materials Engineering (LAGIM), University of 8 May 1945, Guelma, P. O. Box 401, Guelma 24000, Algeria

[5]Department of Process Engineering, Faculty of Technology, University of 8 May 1945, Guelma, BP 401, Guelma 24000, Algeria

[6]School of Natural and Applied Sciences, Kampala International University, Kampala, Uganda

[7]Chemistry Department, Faculty of Exact Sciences, University of El Oued, P.O. Box 789, El Oued 39000, Algeria

[8]Nuclear Research Centre of Birine, P.o. Box 180, Ain Oussera, 17200 Djelfa, Algeria

*Corresponding author (e-mail: (kouadri-2013@hotmail.fr)

All across the world, plant-based medicines and therapeutics are receiving increasing attention. Many studies have been performed, with several modern analytical techniques applied to assess the quality of these medicinal plants. Notably, the use of chemometrics in the study of medicinal plants is natural and essential. This study developed an easy and accurate analytical approach for classifying many plants using HPLC, UV-vis spectroscopy, and chemometrics analysis. The study analyzed 90 samples, and the differences and similarities between these were determined using Principal Component Analysis (PCA) and Hierarchical Clustering (HCA) analysis. The results were validated using the Partial Least Squares method (PLS) and Partial Least Squares Discriminant Analysis (PLS-DA). The HCA and PCA successfully distributed the samples into separate groups based on HPLC and UV-vis data, while PLS and PLS-DA analysis confirmed these results (R^2 = 0.849 - 0.891, RMSEE = 8.83 - 4.907, and RMSEcv =11.81 - 7.399, respectively). Bioactive compounds in plants such as flavonoids, phenolic acids, etc. were described and used to classify the plants. Based on these findings, closely-related plant species may be identified and differentiated using this method.

Keywords: Plant medicines; chemometrics analysis; classification; analytical approach; bioactive compounds

Received: July 2022; Accepted: January 2023

Natural products are biologically-active substances produced from natural sources such as plants, animals, and microorganisms. Historically, most new drugs have been generated from plants (secondary metabolites) and compounds derived from natural products [1]. The main compounds of herbal composites are polyphenols (e.g., flavonoids, phenolic acids, stilbenes), glycosides, tannins, diterpenoids, resins, alkaloids, and many more [2]. Natural products from various geographical origins have varying qualities, compositions, and prices, making it necessary to authenticate their quality and composition to ensure efficacy [3]. When drying and processing techniques prevent botanical and traditional microscopic identification, the authenticity of medicinal plants becomes a challenge. Due to their diversity in terms of chemical composition, scarcity, and unpredictability, the study of the active components of medicinal plants also becomes a difficult undertaking. However, the application of modern analytical tools has helped in generating large amounts of data. Researchers studying medicinal plants now have to overcome the challenge of obtaining sufficient chemical data and performing the required analyses. Chemometrics has proven to be the solution to these challenges [3, 4].

Fingerprint analysis and UV spectroscopy are valuable tools for determining the quality of herbal medicines [5] and distinguishing easily confused herbs

[6]. The systematic approach of selecting several major peaks or a specific region of the fingerprint has been accepted [7]. However, due to the complexity of the secondary metabolites found in herbs, this approach has been shown to be insufficient [8, 9]. Consequently, for species authentication, the entire HPLC fingerprint profiles of the samples have to be evaluated [10, 11]. Chemical pattern recognition provides a reliable and objective method for classifying plant samples based on their fingerprint profiles. In combination with chemometrics such as PCA and HCA, chromatography fingerprinting has recently become one of the most widely-used methods for assessing the chemical profiles of various botanicals[13]. UV-visible and infrared spectroscopy, in combination with chemometrics, have previously been shown to be effective tools for food system differentiation. These were successful screening tools in authenticating Yemeni Sidr honey [14, 15] and various thyme and curcuma samples.

The scope of this study includes the application of chemical fingerprints generated by different analytical techniques (HPLC, UV-vis Spectroscopy) combined with chemometrics methods for the quality assessment of medicinal plants and their bioactive components. A comparison between different analytical techniques and different chemometrics methods was made to determine which one was more appropriate to develop an efficient strategy. The choice of fingerprint techniques and chemometrics methods were made based on the study objectives.

This study aims at clarifying the qualitative and quantitative applications of HPLC and UV-vis spectroscopy for the quality assessment of medicinal plants, such as the differentiation of medicinal plants from different geographical origins, botanical species, growth years, harvest time, processing methods, and the recognition of confused and/or adulterated materials. Multivariate calibration analysis was used to quantify the chemical components or adulteration ratio, and regression analysis was used to screen the quality control markers by modeling fingerprint efficiency relationships. Algeria is known for its abundant plant life. The large size of Algeria and its climatic diversity have resulted in a diverse flora [16]. In this study we developed a taxonomic model for 90 plant samples using chemometrics analysis (PCA, HCA, and PLS) combined with HPLC and UV-vis data, that provides a fast, accurate and feasible method for the classification of Algerian plants, along with building a reference database for the screening of multiple species of plants.

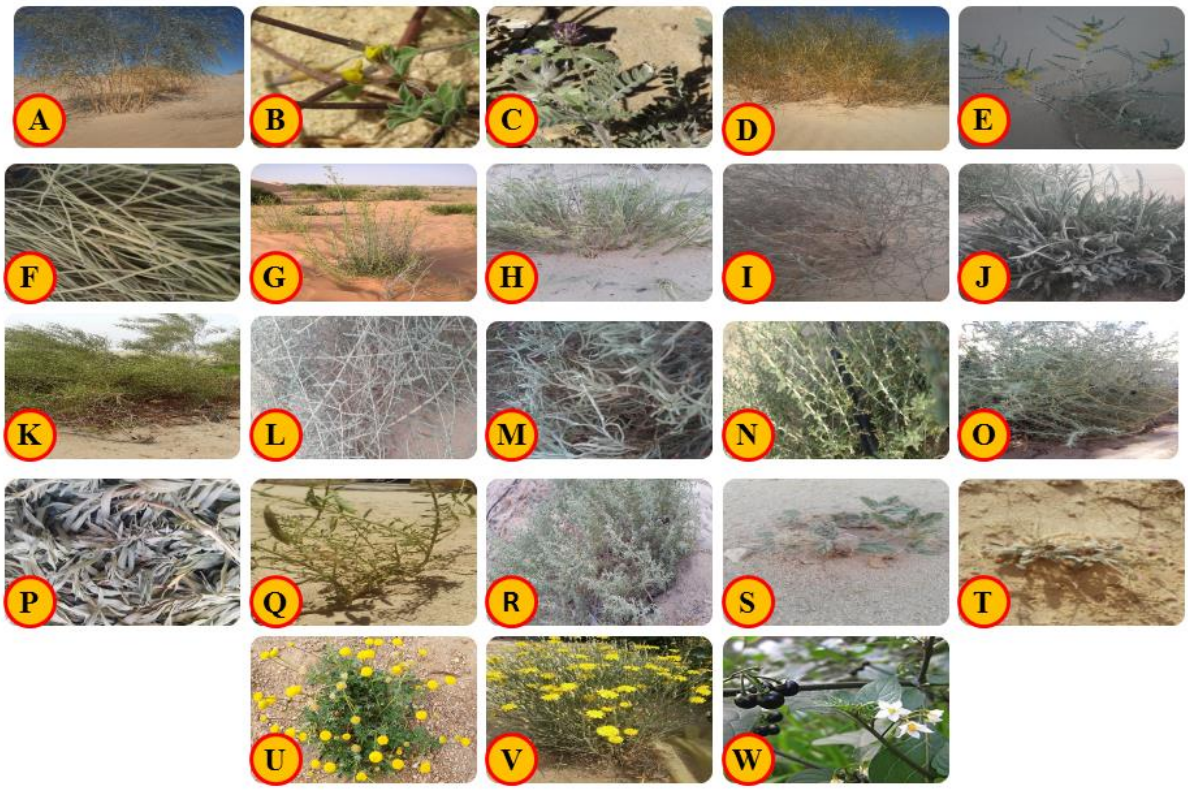


Figure 1. A. *Retama Retam Web b*; **B.** *Lotus halophylus Boiss*; **C.** *Astragalus cruciatus Link*; **D.** *Genista saharae Cosson et Dur.*; **E.** *Astragalus gombiformis Bomel.*; **F.** *Ephedra alata DC.* ; **G.** *Eurphorbia guyoniana Bois et Reut.*; **H.** *Cyperus conglomerates*; **I.** *Heliathemum lipii (L.)*; **J.** *Pers;Plantago albicans L.*; **K.** *Calligonum comosum L'her.*; **L.** *Tamarix boveana* ; **M.** *Limoniastrum guyonianum Dur.*; **N.** *Traganum nudatum Del.*; **O.** *Bassia muricata (L.)*; **P.** *Atriplex halimus L.*; **Q.** *Cleome arabica L*; **R.** *Zygophyllum album L*; **S.** *Neurada procumbens L*; **T.** *Erodium glaucophyllum L'Her.*; **U.** *Launaea resedifolia O. K.*; **V.** *Matricaria pubescens*; **W.** *Solanum nigrum L.*

MATERIALS AND METHODS

Standards and Solvents

The standards and reagents used were of analytical grade. The solvents, i.e., methanol 98%, acetonitrile, and acetic acid, were of HPLC grade, purchased from Sigma-Aldrich, Germany. The reference compounds such as gallic acid (GA), chlorogenic acid (CGA), vanillic acid (VA), caffeic acid (CA), vanillin (V), p-coumaric acid (p-CA), rutin (RU), naringin (NAR), and quercetin (QC), were obtained from Alfa Aesar, United States.

Plant Materials

Ninety samples (thirty-four plants belonging to twenty distinct families) were collected from different regions in El Oued (33°22′06″N and 6°52′03″E), Algeria, between March and May 2019. They were identified by an engineer specializing in botany. The detailed information on the samples can be seen in Table 1 and Figure 1. All the samples were stored under dark, dry and cool conditions at room temperature for a period of one to two months prior to analysis.

Table 1. Sources of samples (Species and Family, The use in traditional medicine) [17].

Sample code	Species	Family	Uses in traditional medicine
1:1	*Astragalus gysensis Del*	Fabaceae (Leguminose)	Analgesic, antiseptic and anti-inflammatory
1:2	*Astragalus gysensis Del*		
1:3	*Astragalus gysensis Del*		
2:4	*Retama Retam Webb*		
2:5	*Retama Retam Webb*		
2:6	*Retama Retam Webb*		
3:7	*Lotus halophylus Boiss*		
3:8	*Lotus halophylus Boiss*		
3:9	*Lotus halophylus Boiss*		
4:10	*Astragalus cruciatus Link*		
4:11	*Astragalus cruciatus Link*		
4:12	*Astragalus cruciatus Link*		
5:13	*Genistasaharae Cosson et Dur*		Colds, respiratory system problems
5:14	*Genistasaharae Cosson et Dur*		
5:15	*Genistasaharae Cosson et Dur*		
6:16	*Astragalus gombiformis Bomel*		Scorpion stings and snake bites
6:17	*Astragalus gombiformisBomel*		
6:18	*Astragalus gombiformisBomel*		
7:19	*Eurphorbia guyoniana Bois and Reut*	Euphorbiaceae	Snake bites
7:20	*Eurphorbia guyoniana Bois and Reut*		
7:21	*Eurphorbia guyoniana Bois and Reut*		
8:22	*Ephedra alata DC*	Ephedraceae	Colds, influenza, respiratory problems, hypertension
8:23	*Ephedra alata DC*		
8:24	*Ephedra alata DC*		
9:25	*Helianthemum lipii (L.) Pers*	Cistaceae	Skin lesions
9:26	*Helianthemum lipii (L.) Pers*		
9:27	*Helianthemum lipii (L.) Pers*		
10:28	*Cyperus conglomeratus*	Cyperaceae	
10:29	*Cyperus conglomeratus*		
10:30	*Cyperus conglomeratus*		
11:31	*Calligonum comosum L'her*	Polygonaceae	Scorpion stings and snake bites
11:32	*Calligonum comosum L'her*		
11:33	*Calligonum comosum L'her*		
12:34	*Plantago albicans L*	Plantaginaceae	/
12:35	*Plantago albicans L*		
12:36	*Plantago albicans L*		
13:37	*Limoniastrumguyonianum Dur*	Plumbaginaceae	Scorpion stings and snake bites, constipation, anemia
13:38	*Limoniastrumguyonianum Dur*		
13:39	*Limoniastrumguyonianum Dur*		
14:40	*Tamarix boveana*	Tamricaceae	Antiseptic, burns, illnesses of the kidney, diarrhea, anemia
14:41	*Tamarix boveana*		
14:42	*Tamarix boveana*		

Table 1. Continued

15:43	*Traganumnudatum Del*	Chenopodiaceae	Rheumatism, skin diseases, against diarrhea
15:44	*Traganumnudatum Del*		
15:45	*Traganumnudatum Del*		
16:46	*Bassia muricata (L.)*		Analgesic, antiseptic and anti-inflammatory
16:47	*Bassia muricata (L.)*		
16:48	*Bassia muricata (L.)*		
17:49	*Atriplex halimus L*		
17:50	*Atriplex halimus L*		
18:51	*Zygophyllum album L*	Zygophyllaceae	Diabetes, purgative and laxative, indigestion, diuretic
18:52	*Zygophyllum album L*		
18:53	*Zygophyllum album L*		
19:54	*Brocchia cinerea Vis*	Asteraceae	Scorpion stings and snake bites, diuretic
19:55	*Brocchia cinerea Vis*		
19:56	*Brocchia cinerea Vis*		
20:57	*Launaea glomerata (Coss.) Hook*		
21:58	*Matricaria pubescens (desf) Schultz*		
21:59	*Matricaria pubescens (desf) Schultz*		
21:60	*Matricaria pubescens (desf) Schultz*		
22:61	*Moltkia ciliata (Forsk.) Maire*	Boraginaceae	/
22:62	*Moltkia ciliata (Forsk.) Maire*		
22:63	*Moltkia ciliata (Forsk.) Maire*		
23:64	*Silene villosa forsk*	Caryophyllaceae	
23:65	*Silene villosa forsk*		
23:66	*Silene villosa forsk*		
24:67	*Atractylis flava L*	Asteraceae	
25:68	*Solanum nigrum L*	Solanaceae	Diuretic, chronic enlargement of liver, dysentery and piles
25:69	*Solanum nigrum L*		
25:70	*Solanum nigrum L*		
26:71	*Malcolmia aegyptiaca Spr*	Brassicaceae	/
26:72	*Malcolmia aegyptiaca Spr*		
26:73	*Malcolmia aegyptiaca Spr*		
27:74	*Matthiola livida DC*		
27:75	*Matthiola livida DC*		
27:76	*Matthiola livida DC*		
28:77	*Launaeare sedifolia O. K*	Asteraceae	/
28:78	*Launaeare sedifolia O. K*		
28:79	*Launaeare sedifolia O. K*		
29:80	*Salsola foetida (sel)*	Chenopodiaceae	/
29:81	*Salsola foetida (sel)*		
29:82	*Salsola foetida (sel)*		
30:83	*Erodium glaucophyllum L'Her*	Geraniaceae	Diarrhea, colds, influenza and problems of respiratory system
30:84	*Erodium glaucophyllum L'Her*		
30:85	*Erodium glaucophyllum L'Her*		
31:86	*Cutandia Dichotoma (forsk.) Trab*	Poaceae	/
31:87	*Cutandia Dichotoma (forsk.) Trab.*		
32:88	*Onopordon macracanthum schousb.*	Asteraceae	
33:89	*Cleome arabica L*	Capparidaceae	Rheumatism, diuretic
34:90	*Neurada procumbens L*	Rosaceae	Analgesic, antiseptic and anti-inflammatory

Preparation of Reference and Sample Solutions

All dry samples were accurately weighed and dissolved in methanol as stock solutions. Each of the stock solutions was diluted to a series of concentrations for the establishment of calibration curves. The preparation conditions of the sample solution, including extraction method, extraction solvent, solid-liquid ratio, and extraction time, were optimized. The final sample solution preparation method was as follows: 200 mg of the dried powdered sample (0.30−0.45 mm) was ultrasonically (40 kHz, 250 w) extracted with 2 mL

methanol at room temperature for 30 min. Then, the methanol extract was filtered, and the filtrate evaporated to dryness in a 30 °C water bath.

To meet the technical requirements of qualitative analysis, the extracts with low concentrations were quantitatively diluted to 10mL with methanol. The obtained solutions were filtered with 0.45 µm nylon membrane filters for further HPLC analysis (Figure 2) [18, 19].

Chromatographic Conditions and Method Validation

For quantitative analysis, HPLC was performed on a Prominence liquid chromatograph, with a thermostatic column compartment, online degasser and a UV-visible detector model SPD-20A. The analytical column used was a Shim-pack VP-ODS C18 (4.6 mm × 250 mm, 5 µm), Shimadzu Co., Japan. The mobile phase was a mixture of acetonitrile and 0.2 % acetic acid in water. The contents of the mobile phase were filtered before use through a 0.45 µm membrane filter, sonicated and pumped from the solvent reservoir to the column HPLC at a flow rate of 1 mL/min, while the eluent was detected at 268 nm. Evaluation of each standard was repeated three times [20].

A linear gradient was used for elution as described below (Table 2). The column temperature was ambient and the volume of injection was 10 µL. Prior to injection, the column was equilibrated with the mobile phase for 40 - 50 min.

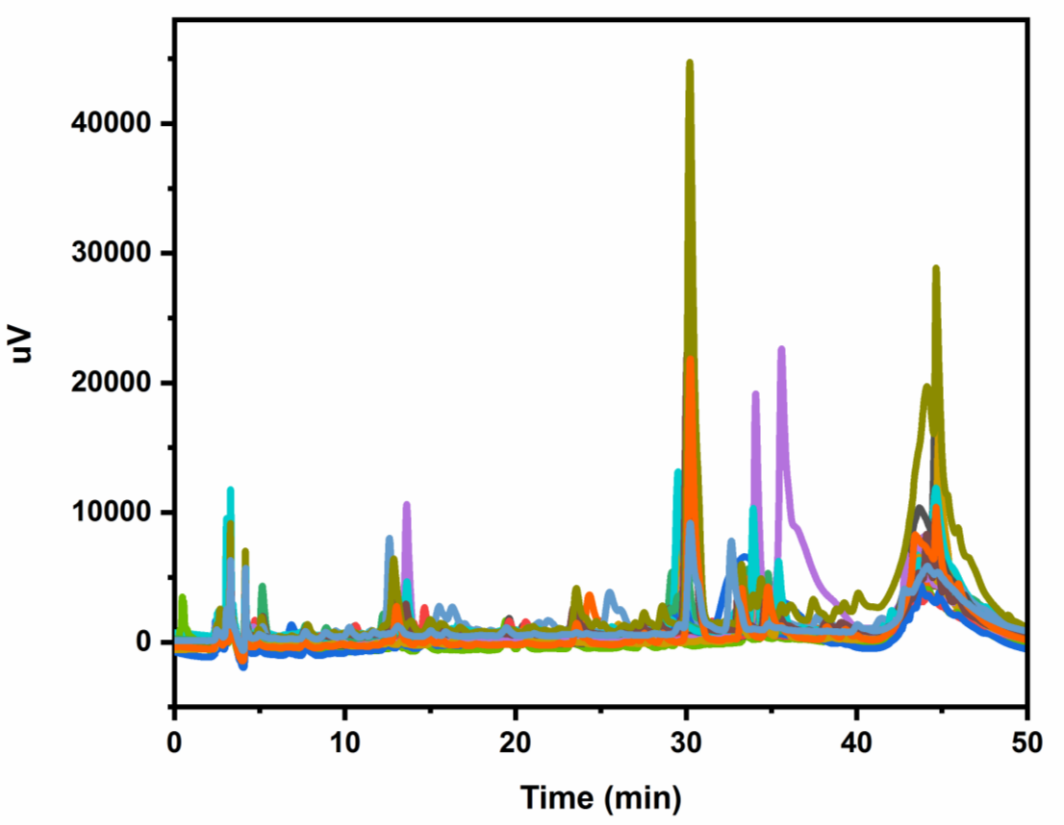


Figure 2. Chromatographic HPLC profiles of the extracted plants.

Table 2. Gradient programme for elution of phenolic acids.

Time (min)	CH_3COOH %	CH_3CN %
0	90	10
6	86	14
16	83	17
23	81	19
28	77	23
35	60	40
38	90	10
50	90	10

UV Spectroscopy

Preparation of Sample Solution

200 mg of the dried powdered sample (0.30 − 0.45 mm) was ultrasonically (40 kHz, 200 w) extracted with 2 mL methanol at room temperature for 30 min. The extract was filtered, and the filtrate was diluted 80 times with methanol before UV–vis analysis [21].

Instrument and Spectra Collection

The UV spectra of all samples were collected in the range of 200 – 900 nm using a UV–vis spectrometer (UV-1800 SHIMADZU). Samples were scanned in a 1 cm thick quartz cell (sampling interval of 0.5 nm) with methanol as reference blank at room temperature. Prior to each scan, baseline correction and wavelength accuracy checks were conducted. All samples were measured in triplicate and the mean spectrum was calculated and used (Figure 3) [21].

Chemometric Analysis

HPLC and UV-vis Spectral Data Pre-treatment

The HPLC data sets (6003 x 90 datasets) obtained from Shimadzu HPLC solution software for normalizing and smoothing HPLC chromatogram data were saved digitally. The numerical values from the spectral files were converted to two data sets (rows: samples; columns: retention time) and manually copied to Microsoft Excel 2010. The data were then aligned in rows for samples and columns for retention time, with the noise range from 0 - 50 minutes cut off. MKS Umetrics AB SIMCA 14.1 Software was used to perform PCA, HCA, PLS, and PLS-DA on all HPLC sample data.

The HPLC chromatogram data was converted to an excel sheet for multivariate analysis, and different chromatograms and chromatogram areas were chosen. To reduce the effect of the baseline and systematic noise, pre-treatment of chromatogram data is a critical step before chemometrics analysis. Baseline correction (preprocessed chromatogram) was used in this study to reduce scatter effects from chromatograms by centering and scaling each one.

UV data sets (1403 x 90 datasets) of standardized and smoothed data from Shimadzu UV solution software were saved digitally. The numerical values from the spectral files were converted to two data sets (rows: samples; columns: wavelength) and manually copied to Microsoft Excel 2010. The data were then aligned in rows for samples and columns for wavelength, with the noise range of 200 - 900 nm cut out. MKS Umetrics AB SIMCA 14.1 Software was used to perform PCA, HCA, PLS, and PLS-DA on all UV sample data.

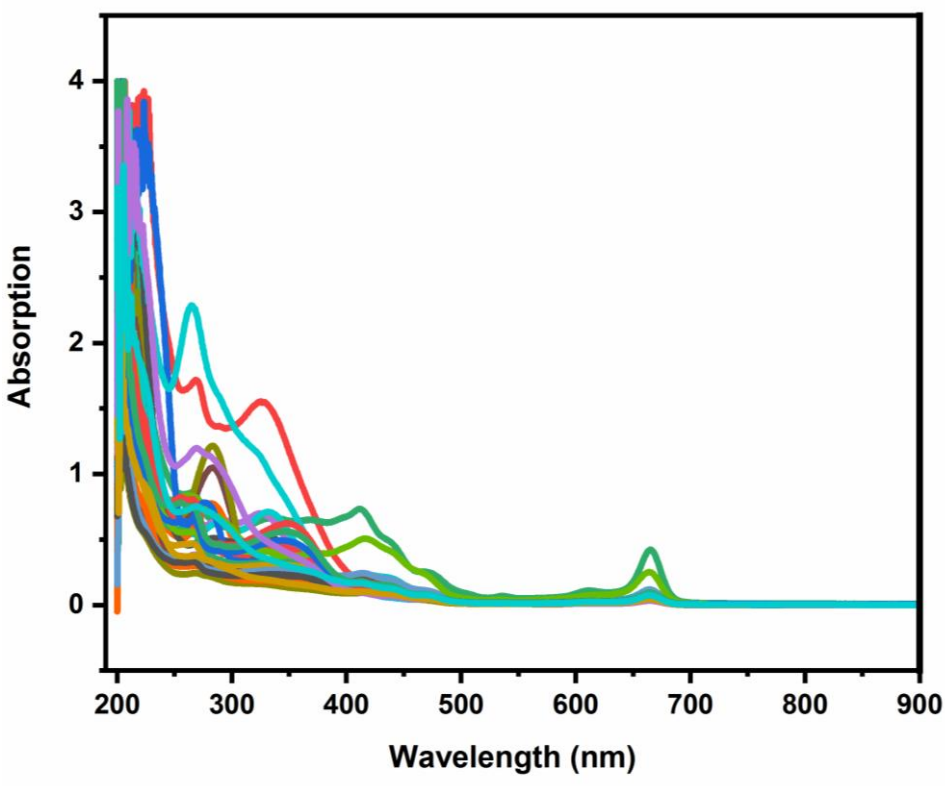


Figure 3. The UV-vis profiles of the extracted plants species.

UV-vis spectral data was converted into an Excel sheet for multivariate analysis, and spectra and distinct spectral regions were chosen. To lessen the effect of light dispersion, baseline variance, and systematic noise, spectral data were pre-treated before chemometrics analysis. Pretreatment with baseline correction (preprocessed spectra) was used to reduce scatter effects from spectra by centering and scaling each spectrum in this investigation.

Principal Component Analysis (PCA)

PCA is a technique for identifying latent variables, or linear combinations of the original variables. The latent variables optimize the description of the remaining variance in the data matrix. These hidden variables are referred to as Principal Components (PCs). The first PC describes the greatest variation. The projections of fingerprints on orthogonal PCs are referred to as scores [22, 23].

Hierarchical Cluster Analysis (HCA)

HCA is a clustering technique that examines how samples are divided into groups and within groups in a hierarchical structure [24]. The results of HCA are commonly shown as a dendrogram, which is a tree-like diagram that shows how samples are organized and related [25].

Partial Least Squares (PLS)

The PLS methodology is a well-known multivariate calibration method. The PLS model is constructed using a set of new orthogonal variables (often referred to as PLS factors) that optimize the covariance between the PLS factors and the response variable (s) [26]. PLS can correlate and simultaneously decompose the measure matrix X and the response matrix Y by examining their relationship. PLS finds the optimum linear correction model by projecting X and Y to new factor spaces. It is especially effective for datasets with many correlations and small sample sizes [5, 27].

Partial Least Squares Discriminant Analysis (PLS-DA)

The discriminant method PLS-DA is derived from PLS regression models [21]. PLS-DA is a focused metabolomics approach that detects marker molecules in defined groups of samples that show variation. The links between samples and the variables that define group separation are depicted in score and loading charts [28].

RESULTS AND DISCUSSION

PCA and HCA

PCA and HCA were carried out on the 90 samples by employing integrated UV-vis spectra (from 200 - 900 nm); the results are shown as score plots to each pre-processing method used (Figure 4a). Using at least 20 plants families (Table 1) in this study, along with the HCA classification as shown (Figure 4b), we divided them into four groups. Consequently, the plants studied here possessed different phenolic and flavonoid components, thus it is very important to determine which species were related in regards to their chemical composition. This partly depends upon the richness of the medicinal plants in terms of polyphenolic compounds, especially catechins, flavonoids and phenolic acids; this was confirmed by what was simply classified using PCA, where the results were good and the divisions were clear and interconnected based on the concentration of polyphenols in different plants.

Groups of comparable patterns were obtained for the samples and analyzed by charting the first two PCs in a PCA plot. The first two PCs are typically utilized because they are the most useful components for analysis and also because they represent the majority of the data variance. In this study, PCA was applied to UV-Vis spectra in the 200 - 900 nm range to evaluate the possibility of classifying all samples. PCA was applied to all 90 objects x 1403 variables data matrices. Figure 4a displays the PCA score plot utilizing the first two PCs, which accounts for 84 % of the total, with PC1 accounting for around 87 percent and PC2 for about 3%. We can observe from the PCA plot that all of the samples studied were classified into four groups [29].

HCA was used to identify the samples based on their UV-vis spectral data similarities and differences. Utilizing UV-vis spectral data and a similarity threshold of 0.5, an average group-linking method based on Pearson's correlation was used to create a hierarchical agglomerative dendrogram. The clustering dendrogram (Figure 4b) demonstrated that all of the samples were grouped into four main clusters, confirming the PCA results.

The results of the unsupervised pattern recognition techniques (HCA and PCA) clarified the separation of the samples into discriminate clusters based on their geographical origin and content of phenolic and flavonoid components, which reflected the significant variation in their composition. As a result, more precise isolation of plant samples from different Algerian regions is required. Due to this, supervised pattern recognition techniques such as PLS and PLS-DA (Partial Least Square Discriminant Analysis) were used to verify the results produced from the unsupervised techniques [30].

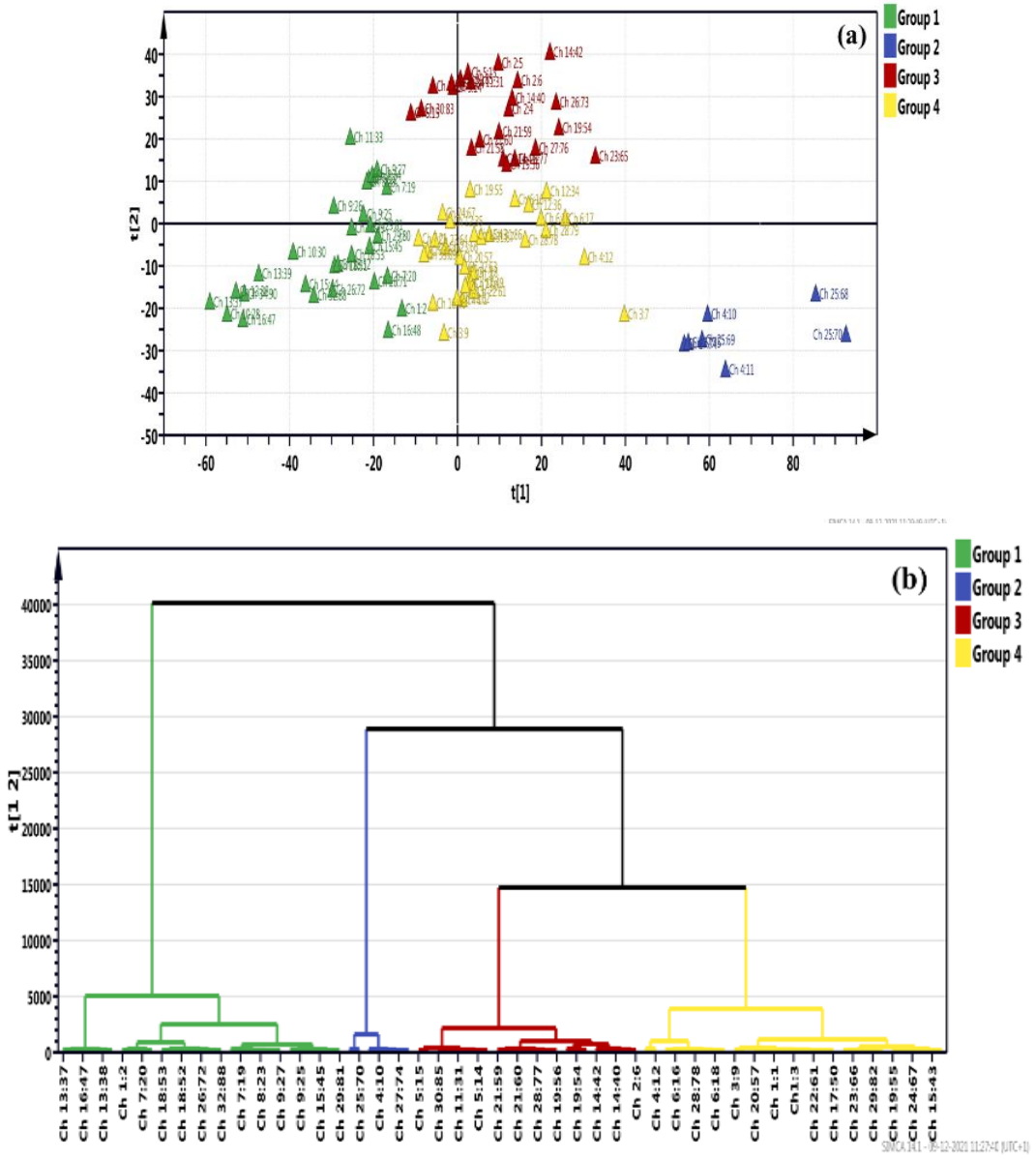


Figure 4. The UV-vis data matrix for medicinal plant samples (a) PCA score plot, (b) HCA-PCA score plot.

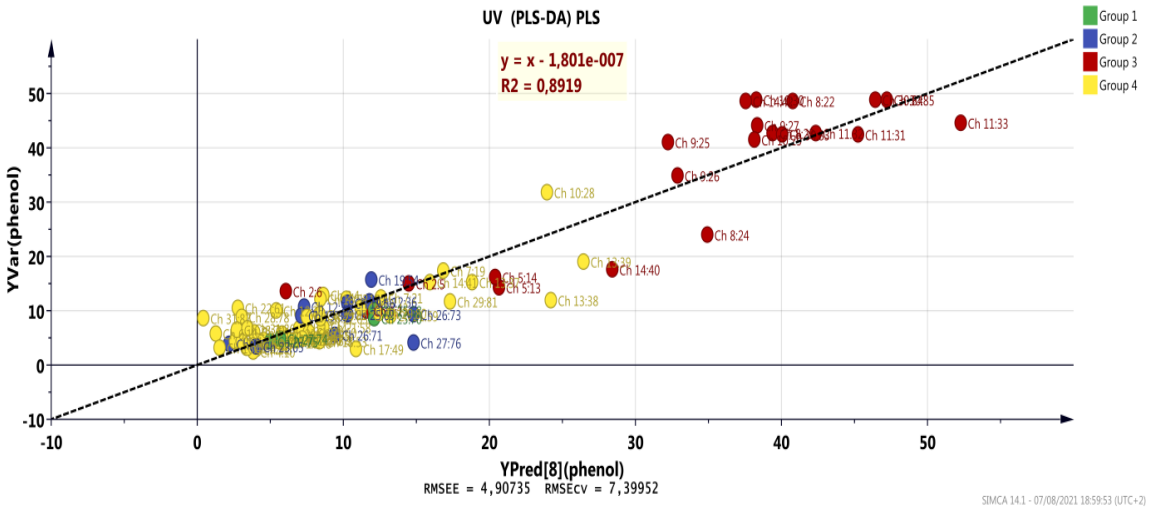


Figure 5. Partial Least Squares (PLS) plot derived from UV-visible spectra of medicinal plants.

As indicated in Figure 5, the chemometrics model utilized in this study was Partial Least Squares (PLS and PLS-DA). A comprehensive cross-validation classification model was used to apply the UV-Visible spectra of 90 samples. The predicted and reference samples had R-squared values of 0.848 and 0.891, respectively. A calibration model with an R-squared value of more than 0.891 is considered an excellent prediction. The PLS model was adequate to differentiate the content of phenolic and flavonoid components for various plant samples, as seen by the high R-squared value of close to 1.0. Furthermore, the mean squared errors (RMSECV) for both the projected and a reference sample were low enough, at 7.39 and 4.9, respectively. These low values indicate that the PLS model was obverting, which is a favourable finding for this study [31, 32]. The actual and predicted concentration values (mg/g) are given in Table 2.

Table 3. Comparison of the experimental and predicted total amounts of polyphenols and flavonoids.

Sample code	C (mg/g) (polyphenol) Excremental	C (mg/g) (polyphenol) Predict	C (mg/g) (flavonoid) Excremental	C (mg/g) (flavonoid) Predict
Ch 1:1	3,29	3,37	1,62	1,56
Ch 1:2	4,12	5,89	1,01	0,93
Ch1:3	4,27	5,57	1,38	1,40
Ch 2:4	12,79	8,65	1,58	1,42
Ch 2:5	14,94	14,50	1,63	1,66
Ch 2:6	13,61	6,05	2,74	2,00
Ch 3:7	3,90	4,10	1,00	1,76
Ch 3:8	2,84	4,02	0,47	0,90
Ch 3:9	4,61	8,07	0,46	0,80
Ch 4:10	2,45	3,81	1,52	1,40
Ch 4:11	3,80	2,15	2,91	2,33
Ch 4:12	3,08	3,44	0,91	1,51
Ch 5:13	14,26	20,67	2,01	1,76
Ch 5:14	16,32	20,38	1,73	1,93
Ch 5:15	9,96	11,57	1,49	1,86
Ch 6:16	5,69	1,24	3,97	3,26
Ch 6:17	5,98	3,92	2,20	2,31
Ch 6:18	5,65	7,18	4,18	3,82
Ch 7:19	17,30	16,83	1,50	1,20
Ch 7:20	4,82	7,00	0,33	0,68
Ch 7:21	12,47	12,55	0,85	1,34
Ch 8:22	48,64	40,78	1,64	2,07
Ch 8:23	42,65	39,42	0,42	1,06
Ch 8:24	23,91	34,90	1,34	2,18
Ch 9:25	41,06	32,25	2,29	1,32
Ch 9:26	34,90	32,86	0,92	0,53
Ch 9:27	44,12	38,34	0,70	1,87
Ch 10:28	31,92	23,92	0,22	0,27
Ch 10:29	41,63	38,12	0,02	1,07
Ch 10:30	48,81	38,27	1,50	0,83
Ch 11:31	42,48	45,22	3,62	2,43
Ch 11.32	42,69	42,35	0,98	1,18
Ch 11:33	44,52	52,29	2,07	0,94
Ch 12:34	10,75	7,30	0,84	0,60
Ch 12:35	9,963	5,42	0,65	0,50
Ch 12:36	11,81	11,80	0,26	0,37

Table 3. Continued

Ch 13:37	15,24	18,80	0,44	-0,37
Ch 13:38	11,86	24,21	0,35	-0,33
Ch 13:39	19,10	26,45	1,00	0,25
Ch 14:40	17,55	28,44	2,78	2,57
Ch 14:41	15,36	15,95	3,98	2,59
Ch 14:42	48,73	37,56	4,25	3,82
Ch 15:43	10,49	11,20	0,19	0,77
Ch 15:44	6,80	6,98	0,37	0,024
Ch 15:45	12,22	10,20	0,86	1,40
Ch 16:46	4,29	2,54	0,001	0,87
Ch 16:47	3,08	1,53	1,01	0,45
Ch 16:48	4,02	5,90	1,69	0,75
Ch 17:49	2,96	10,84	1,11	0,99
Ch 17:50	4,51	4,56	1,13	1,18
Ch 18:51	4,59	6,68	0,40	1,17
Ch 18:52	12,20	8,40	0,82	0,95
Ch 18:53	4,39	8,39	0,001	1,01
Ch 19:54	15,71	11,90	6,38	4,36
Ch 19:55	9,14	7,09	4,11	2,54
Ch 19:56	11,51	10,29	3,32	2,62
Ch 20:57	5,94	5,18	1,19	0,55
Ch 21:58	7,08	8,63	0,56	0,94
Ch 21:59	9,18	13,21	1,33	1,81
Ch 21:60	6,02	7,65	0,79	1,01
Ch 22:61	10,63	2,79	0,59	1,36
Ch 22:62	5,29	5,97	2,59	1,67
Ch 22:63	8,71	8,43	0,61	1,42
Ch 23:64	4,35	3,04	1,28	1,47
Ch 23:65	3,49	4,03	0,72	3,13
Ch 23:66	5,78	5,80	1,38	1,63
Ch 24:67	3,76	7,35	0,19	1,64
Ch 25:68	9,98	12,11	5,20	5,13
Ch 25:69	9,41	10,26	4,34	3,87
Ch 25:70	8,73	12,11	6,19	6,22
Ch 26:71	5,45	9,43	0,95	1,16
Ch 26:72	4,82	7,21	0,75	1,12
Ch 26:73	9,24	14,79	2,93	3,22
Ch 27:74	4,88	5,66	2,78	2,41
Ch 27:75	4,37	5,035	1,52	1,98
Ch 27:76	4,16	14,83	1,24	3,32
Ch 28:77	8,94	7,48	0,34	-0,36
Ch 28:78	8,82	3,03	2,73	1,45
Ch 28:79	6,53	2,71	0,53	1,38
Ch 29:80	9,65	12,42	0,53	-0,22
Ch 29:81	11,71	17,33	0,03	0,04
Ch 29:82	9,31	8,39	0,14	0,33
Ch 30:83	42,52	40,02	0,84	1,41
Ch 30:84	48,85	46,41	1,09	1,67
Ch 30:85	48,93	47,24	0,94	1,69

Table 3. Continued

Ch 31:86	6,80	3,46	1,22	0,81
Ch 31:87	8,55	0,40	1,35	0,91
Ch 32:88	8,92	9,71	0,06	-0,18
Ch 33:89	3,74	8,12	0,94	2,01
Ch 34:90	22,26	34,29	0,09	-0,19

HPLC Method of Quantitative Analysis

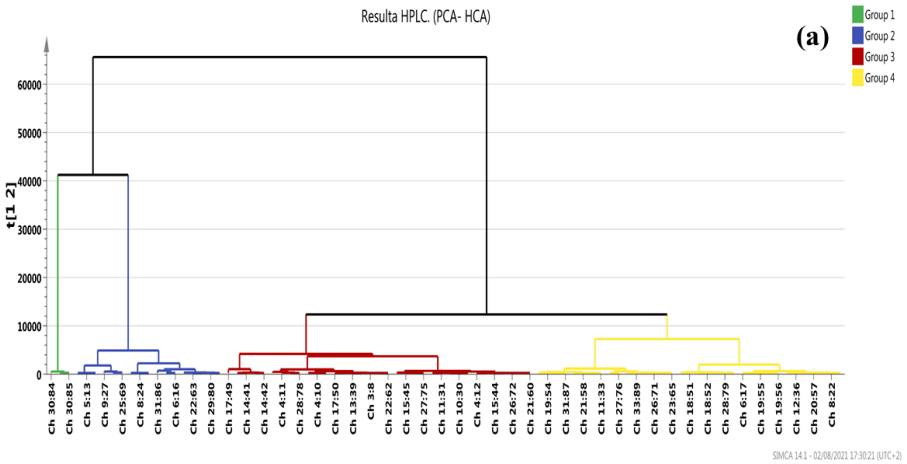
Determination of the macroconstituents of medicinal plants, such as phenolic and flavonoid compounds, total carbohydrates, amino acids, and polysaccharides, is critical for maintaining the quality of these herbs and herbal products. However, the importance of such components in the bioactive quality analysis of their extracts should be considered, because macro-constituents and polyphenolic compounds have a significant impact on the extracts' total biological activity [33, 34].

Table 3 shows the findings of preliminary spectrophotometric experiments for screening plant species based on their quantitative content of fundamental (phenolic and flavonoid) macroconstituents. This demonstrates the significant variation in the components of the medicinal plants. The natural clustering of samples according to their fingerprints was determined using HCA and PCA. In this scenario, HCA was used to examine data from HPLC chromatograms of 90 samples (Figure 2). The software was given a 90-row, 6003-column data matrix to work with. For the depiction of the data matrix, scatter plots were defined by the relationship between each principal component (PC). The goal of this analysis was to extract the under-lying information from multivariate original data by converting and reducing the dimensions of the original data matrix for samples and variables into the product of two matrices. Sample information was presented as scores, while loadings focused on the variables that had the most influence over the difference between groups of samples [35].

For a potential scale-up application of the extracts, a thorough statistical analysis can provide a quick and reliable analytical tool, as shown in many studies [36-38]. To create clusters, clustering analysis used the between-groups linkage method as the amalgamation rule and the cosine method as the metric. Phenolic acid concentration was thus an essential metric not only in terms of nutritional and dietary value, but also in terms of medicinal and pharmacological importance [36]. Figure 6 shows the findings of the score and loading scatter plots based on the differences in their HPLC fingerprints. The samples could be divided into four groups based on the scatter points. The samples in the second and third groups were tightly grouped, but the samples in the first group were widely dispersed. The PC1 versus PC2 biplot in Figure 8 accounted for 70% of the variation, whereas PC2 accounted for 20%. The PCA result was in good agreement with the HCA, which revealed minor variances in the primary contents and differences in the distribution of the main components. The further a variable is situated from the origin, the greater its contribution to the PCA model.

The total concentration (Table 4) in elderberries was determined in this study, and the results are summarized in Table 4 and shown in Figure 6. However, some of the discrepancies in phenolic component amounts were most likely related to the different plant extract composites [39, 40].

The results of the unsupervised pattern recognition techniques (HCA and PCA) explained the sample differentiation into classified clusters based on their chemical components. The HCA procedure was able to find natural clusters of samples based on their fingerprints, and all of the samples were clearly divided into different clusters.



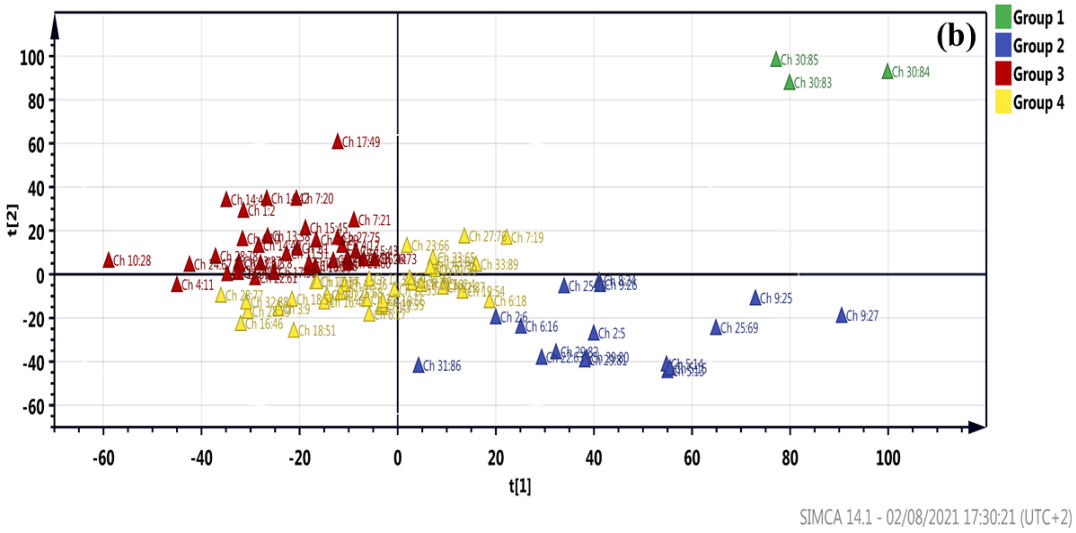


Figure 6. The score plot for the 90 samples: (a) HCA, (b) PCA.

Simultaneously, HPLC chromatogram data provided a powerful classification tool for plant samples, where key component groupings and clusters were found, while the least-squares model clearly demonstrated diverse groupings. A determination coefficient (R^2) was used to determine the regression equation. For each model, R^2 versus RMSEcv values were evaluated, and R^2 was chosen for model robustness with low RMSEE and greatest R^2 values (Figure 7). The following findings show the linear regression coefficients achieved for the best PLS and PLS-DA model outcomes: R^2 =0.849, Q=0.741, RMSEE = 8.83, and RMSEcv=11.81; RMSEE = 8.83, and RMSEcv=11.81 [35].

The results in terms of identifying and distinguishing these samples were satisfactory, indicating that chromatographic fingerprinting aided by pattern recognition could be a viable strategy for plant authentication and classification.

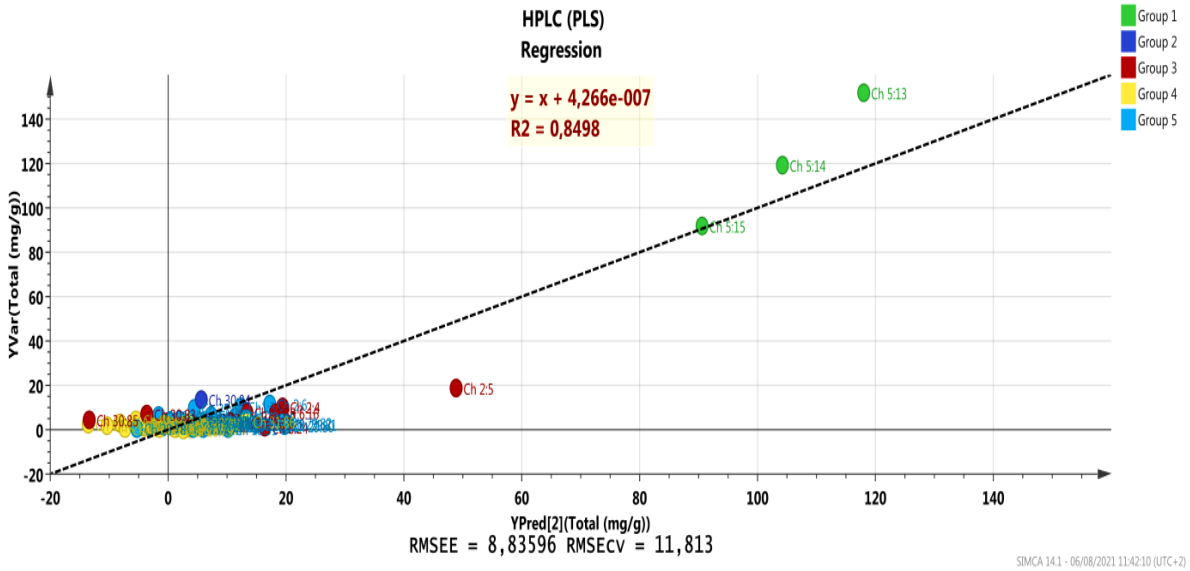


Figure 7. Predicted vs. experimental data (CI 95%) for parameters with the best derived correlations with the HPLC spectra.

Table 4. Comparison of experimental and predicted total concentrations.

Sample code	Species	Family	Total Concentration (mg/g), Experimental	Total Concentration (mg/g), Predicted
Ch 1:1	*Astragalus gysensis Del.*	Fabaceae	3.12	0.14
Ch 1:2	*Astragalus gysensis Del.*	Fabaceae	1.56	-4.57
Ch1:3	*Astragalus gysensis Del.*	Fabaceae	10.44	12.3
Ch 2:4	*Retama Retam Web b.*	Fabaceae	10.23	19.4021
Ch 2:5	*Retama Retam Web b.*	Fabaceae	18.59	48.86
Ch 2:6	*Retama Retam Web b.*	Fabaceae	11.63	17.25
Ch 3:7	*Lotus halophylus Boiss.*	Fabaceae	0.73	5.87
Ch 3:8	*Lotus halophylus Boiss.*	Fabaceae	2.92	-2.57
Ch 3:9	*Lotus halophylus Boiss.*	Fabaceae	6.51	7.25
Ch 4:10	*Astragalus cruciatus Link.*	Fabaceae	3.48	-1.35
Ch 4:11	*Astragalus cruciatus Link.*	Fabaceae	0.88	-2.62
Ch 4:12	*Astragalus cruciatus Link.*	Fabaceae	2.42	0.04
Ch 5:13	*Genista saharae Cosson et Dur.*	Fabaceae	151.69	117.95
Ch 5:14	*Genista saharae Cosson et Dur.*	Fabaceae	119.226	104.135
Ch 5:15	*Genista saharae Cosson et Dur.*	Fabaceae	91.95	90.61
Ch 6:16	*Astragalus gombiformis Bomel.*	Fabaceae	7.82	18.18
Ch 6:17	*Astragalus gombiformis Bomel.*	Fabaceae	3.39	7.78
Ch 6:18	*Astragalus gombiformis Bomel.*	Fabaceae	5.87	11.31
Ch 7:19	*Eurphorbia guyoniana Bois et Reut.*	Euphorbiaceae	3.91	0.005
Ch 7:20	*Eurphorbia guyoniana Bois et Reut.*	Euphorbiaceae	1.75	-10.36
Ch 7:21	*Eurphorbia guyoniana Bois et Reut.*	Euphorbiaceae	1.92	-4.46
Ch 8:22	*Ephedra alata DC.*	Ephedraceae	1.50	2.94
Ch 8:23	*Ephedra alata DC.*	Ephedraceae	1.39	3.22
Ch 8:24	*Ephedra alata DC.*	Ephedraceae	1.47	16.44
Ch 9:25	*Heliathemum lipii (L.) Pers.*	Cistaceae	3.60	10.53
Ch 9:26	*Heliathemum lipii (L.) Pers.*	Cistaceae	2.93	1.24
Ch 9:27	*Heliathemum lipii (L.) Pers.*	Cistaceae	5.37	4.71
Ch 10:28	*Cyperus conglomeratus*	Cyperaceae	6.45	-1.58
Ch 10:29	*Cyperus conglomeratus*	Cyperaceae	2.28	-0.17
Ch 10:30	*Cyperus conglomeratus*	Cyperaceae	3.51	0.39
Ch 11:31	*Calligonum comosum L'her.*	Polygonaceae	3.58	-3.31
Ch 11.32	*Calligonum comosum L'her.*	Polygonaceae	2.82	9.31
Ch 11:33	*Calligonum comosum L'her.*	Polygonaceae	7.4054	13.41
Ch 12:34	*Plantago albicans L.*	Plantaginaceae	1.97	4.33
Ch 12:35	*Plantago albicans L.*	Plantaginaceae	3.48	5.29
Ch 12:36	*Plantago albicans L.*	Plantaginaceae	4.29	1.94
Ch 13:37	*Limoniastrum guyonianum Dur.*	Plumbaginaceae	1.78	-5.07
Ch 13:38	*Limoniastrum guyonianum Dur.*	Plumbaginaceae	2.97	-8.11
Ch 13:39	*Limoniastrum guyonianum Dur.*	Plumbaginaceae	0.88	-4.38
Ch 14:40	*Tamarix boveana*	Tamricaceae	0.82	-3.85
Ch 14:41	*Tamarix boveana*	Tamricaceae	4.28	-5.49
Ch 14:42	*Tamarix boveana*	Tamricaceae	1.726	-2.99
Ch 15:43	*Traganum nudatum Del.*	Chenopodiaceae	1.81	0.65
Ch 15:44	*Traganum nudatum Del.*	Chenopodiaceae	1.15	-0.91
Ch 15:45	*Traganum nudatum Del.*	Chenopodiaceae	0.66	-7.25
Ch 16:46	*Bassia muricata (L.)*	Chenopodiaceae	1.11	5.05

Ch 16:47	*Bassia muricata (L.)*	Chenopodiaceae	0.56	3.82
Ch 16:48	*Bassia muricata (L.)*	Chenopodiaceae	0.86	-2.28
Ch 17:49	*Atriplex halimus L.*	Chenopodiaceae	2.41	-13.61
Ch 17:50	*Atriplex halimus L.*	Chenopodiaceae	1.45	-2.64
Ch 18:51	*Zygophyllum album L.*	Zygophyllaceae	0.81	10.22
Ch 18:52	*Zygophyllum album L.*	Zygophyllaceae	0.43	1.26
Ch 18:53	*Zygophyllum album L.*	Zygophyllaceae	2.33	-0.21
Ch 19:54	*Brocchia cinerea Vis.*	Asteraceae	2.08	7.15
Ch 19:55	*Brocchia cinerea Vis.*	Asteraceae	3.34	12.6
Ch 19:56	*Brocchia cinerea Vis.*	Asteraceae	2.85	4.18
Ch 20:57	*Launaea glomerata (Coss.) Hook.*	Asteraceae	2.91	-1.57
Ch 21:58	*Matricaria pubescens (desf) Schultz.*	Asteraceae	4.97	13.25
Ch 21:59	*Matricaria pubescens (desf) Schultz.*	Asteraceae	3.33	9.02
Ch 21:60	*Matricaria pubescens (desf) Schultz.*	Asteraceae	2.70	4.10
Ch 22:61	*Moltkia ciliata (Forsk.) Maire.*	Boraginaceae	2.49	-0.61
Ch 22:62	*Moltkia ciliata (Forsk.) Maire.*	Boraginaceae	2.20	-3.54
Ch 22:63	*Moltkia ciliata (Forsk.) Maire.*	Boraginaceae	2.57	14.90
Ch 23:64	*Silene villosa forsk.*	Caryophyllaceae	3.33	6.72
Ch 23:65	*Silene villosa forsk.*	Caryophyllaceae	1.53	6.90
Ch 23:66	*Silene villosa forsk.*	Caryophyllaceae	3.95	3.93
Ch 24:67	*Atractylis fhava L.*	Asteraceae	0.88	0.83
Ch 25:68	*Solanum nigrum L.*	Solanaceae	9.45	4.43
Ch 25:69	*Solanum nigrum L.*	Solanaceae	2.25	10.58
Ch 25:70	*Solanum nigrum L.*	Solanaceae	1.84	7.98
Ch 26:71	*Malcolmia aegyptiaca Spr.*	Brassicaceae	2.94	3.36
Ch 26:72	*Malcolmia aegyptiaca Spr.*	Brassicaceae	0.42	-1.42
Ch 26:73	*Malcolmia aegyptiaca Spr.*	Brassicaceae	0.31	-1.49
Ch 27:74	*Mathiola livida DC.*	Brassicaceae	0.81	4.29
Ch 27:75	*Mathiola livida DC.*	Brassicaceae	3.015	-4.06
Ch 27:76	*Mathiola livida DC.*	Brassicaceae	1.38	1.97
Ch 28:77	*Launaea resedifolia O. K.*	Asteraceae	0.81	1.06
Ch 28:78	*Launaea resedifolia O. K.*	Asteraceae	0.57	-5.30
Ch 28:79	*Launaea resedifolia O. K.*	Asteraceae	0.17	2.55
Ch 29:80	*Salsola foetida (sel).*	Chenopodiaceae	1.97	19.76
Ch 29:81	*Salsola foetida (sel).*	Chenopodiaceae	2.38	20.23
Ch 29:82	*Salsola foetida (sel).*	Chenopodiaceae	3.13	19.33
Ch 30:83	*Erodium glaucophyllum L'Her.*	Geraniaceae	7.11	-3.57
Ch 30:84	*Erodium glaucophyllum L'Her.*	Geraniaceae	13.85	5.59
Ch 30:85	*Erodium glaucophyllum L'Her.*	Geraniaceae	4.21	-13.38
Ch 31:86	*Cutandia Dichotoma (forsk.) Trab.*	Poaceae	3.50	13.21
Ch 31:87	*Cutandia Dichotoma (forsk.) Trab.*	Poaceae	3.64	5.63
Ch 32:88	*Onopordon macracanthum schousb.*	Asteraceae	1.76	0.26
Ch 33:89	*Cleome arabica L.*	Capparidaceae	1.34	7.23
Ch 34:90	*Neurada procumbens L.*	Rosaceae	2.11	9.44

CONCLUSION

In the present study, UV-vis spectroscopy and HPLC were applied to the screening and quantification analysis of diverse medicinal plant species using chemometrics. In this work, we first developed a HPLC method for the simultaneous determination of nine components in different plant species. The results showed that the contents of the nine investigated components varied greatly between different species. The PCA and HCA analysis of the HPLC data showed a clear separation between samples. It also provided excellent separation between different species of medicinal plants based on their UV-vis spectra. A supervised pattern recognition

approach (PLS) using the HPLC data or the UV-vis spectral data for distinguishing the various plant samples was also studied. The predicted and reference samples had R-squared values of 0.848 and 0.891. A calibration model with an R-squared value of more than 0.891 was considered an excellent prediction. The PLS model was adequate to differentiate the content of phenolic and flavonoid components in various plant samples, as seen by the high R-squared value of close to 1.0. Furthermore, the mean squared errors (RMSECV) for both the projected and reference samples were low enough, at 7.39 and 4.9, respectively. Compared to traditional identification methods and classical chromatographic techniques, the UV-vis qualitative and quantitative models gave a more rapid and simple approach for the species identification of different plants, and predicted the main active compounds in plants. The method could be a fast and useful procedure for routine quality evaluation and quality control of plants in the manufacturing process and clinical use.

ACKNOWLEDGEMENTS

The authors wish to thank Mr. Ali Tliba and Mr. Sadok Nani from Laboratory Valorization and Technology of Saharan Resources (VTRS), University of El Oued, staff.

REFERENCES

1. Bernardini, S., Tiezzi, A., Laghezza Masci, V. and Ovidi, E. (2018) Natural products for human health: an historical overview of the drug discovery approaches. *Natural product research*, **32(16)**, 1926–1950.

2. Simion, I. M., Casoni, D. and Sârbu, C. (2019) Classification of Romanian medicinal plant extracts according to the therapeutic effects using thin layer chromatography and robust chemometrics. *Journal of Pharmaceutical and Biomedical Analysis*, **163,** 137–143.

3. Yudthavorasit, S., W–, K. and Leepipatpiboon, N. (2014) Characteristic fingerprint based on gingerol derivative analysis for discrimination of ginger (Zingiber officinale) according to geographical origin using HPLC-DAD combined with chemometrics. *Food chemistry*, **158**, 101–111.

4. Du, J., Wan, D., Huang, L., Chen, S. and Qin, M. (2011) Application of chemometrics in quality evaluation of medicinal plants. *Journal of Medicinal Plants Research*, **5(17)**, 4001–4008.

5. Liu, X., Jiang, W., Su, M., Sun, Y., Liu, H., Nie, L. and Zang, H. (2020) Quality evaluation of traditional Chinese medicines based on fingerprinting. *Journal of separation science*, **43(1)**, 6–17.

6. Wang, L., Wang, X. and Kong, L. (2012) Automatic authentication and distinction of Epimedium koreanum and Epimedium wushanense with HPLC fingerprint analysis assisted by pattern recognition techniques. *Biochemical Systematics and Ecology*, **40**, 138–145.

7. Liang, Y. Z., Xie, P. S. and Chan, K. (2010) Perspective of chemical fingerprinting of Chinese herbs. *Planta medica*, **76(17)**, 1997–2003.

8. Wianowska, D. and Gil, M. (2019) Critical approach to PLE technique application in the analysis of secondary metabolites in plants. *TrAC Trends in Analytical Chemistry*, **114**, 314–325.

9. Croteau, R., Kutchan, T. M. and Lewis, N. G. (2000) Natural products (secondary metabolites). *Biochemistry and molecular biology of plants*, **24**, 1250–1319.

10. Tian, R. T., Xie, P. S. and Liu, H. P. (2009) Evaluation of traditional Chinese herbal medicine: Chaihu (Bupleuri Radix) by both high-performance liquid chromatographic and high-performance thin-layer chromatographic fingerprint and chemometric analysis. *Journal of Chromatography A*, **1216(11)**, 2150–2155.

11. Xie, P., Chen, S., Liang, Y. Z., Wang, X., Tian, R. and Upton, R. (2006) Chromatographic fingerprint analysis-a rational approach for quality assessment of traditional Chinese herbal medicine. *Journal of chromatography A*, **1112(1-2)**, 171–180.

12. Shen, Y., Yao, C. L. and Guo, D. A. (2020) Quality assessment of herbal medicines based on chemical fingerprints combined with chemometrics approach: A review. *Journal of Pharmaceutical and Biomedical Analysis*, **185**, 113215.

13. Yi, T., Chen, Q., He, X., So, S., Lo, Y., Fan, L. and Chen, H. (2013) Chemical quantification and antioxidant assay of four active components in Ficus hirtaroot using UPLC-PAD-MS fingerprinting combined with cluster analysis. *Chemistry Central Journal*, **7(1)**, 1–9.

14. Roshan, A. R. A., Gad, H. A., El-Ahmady, S. H., Khanbash, M. S., Abou-Shoer, M. I. and Al-Azizi, M. M. (2013) Authentication of monofloral yemeni sidr honey using ultraviolet spectroscopy and chemometric analysis. *Journal of agricultural and food chemistry*, **61(32)**, 7722–7729.

15. Beltran, V., Salvadó, N., Butí, S. and Cinque, G. (2015) Micro infrared spectroscopy discrimination capability of compounds in complex matrices of thin layers in real sample coatings from artworks. *Microchemical journal*, **118**, 115–123.

16. Hamza, N., Berke, B., Umar, A., Cheze, C., Gin, H. and Moore, N. (2019) A review of Algerian medicinal plants used in the treatment of diabetes. *Journal of ethnopharmacology*, **238**, 111841.

17. Jannah, R., Rafi, M., Heryanto, R., Kautsar, A. and Septaningsih, D. A. (2018) UV-Vis spectroscopy and chemometrics as a tool for identification and discrimination of four Curcuma species. *International Food Research Journal*, **25(2)**.

18. Aboulwafa, M. M., Youssef, F. S., Gad, H. A., Sarker, S. D., Nahar, L., Al-Azizi, M. M. and Ashour, M. L. (2019) Authentication and discrimination of green tea samples using UV–vis, FTIR and HPLC techniques coupled with chemometrics analysis. *Journal of Pharmaceutical and Biomedical Analysis*, **164**, 653–658.

19. Sirisomboon, P. and Posom, J. (2019) On-line measurement of activation energy of ground bamboo using near infrared spectroscopy. *Renewable Energy*, **133**, 480–488.

20. Gusti, N., Oktarina, D., Elvia, R., Nursa'adah, E., Wardhana, R. W., Sudaryono, A. and Firdaus, M. L. (2021) Facile Detection of Oil Adulteration using UV-Visible Spectroscopy Coupled with Chemometrics Analysis. *Science and Technology Indonesia*, **6(1)**, 14–18.

21. Qureshi, M. N., Stecher, G., Sultana, T., Abel, G., Popp, M. and Bonn, G. K. (2011) Determination of carbohydrates in medicinal plants ‐ comparison between TLC, mf ‐ MELDI ‐ MS and GC ‐ MS. *Phytochemical Analysis*, **22(4)**, 296–302.

22. Qureshi, M. N., Stecher, G. and Bonn, G. K. (2013) Matrix free material enhanced laser desorption ionization mass spectrometric analysis of amino acids in Althaea officinalis, Matricaria chamomilla, and Taraxacum officinale. *Analytical Letters*, **46(1)**, 29–34.

23. Maimaiti, Z., Turak, A., Ma, Q. L., Liu, G. and Aisa, H. A. (2020) Quantitative determination of marker compounds and fingerprint analysis of the seeds of Vernonia anthelmintica. *International Journal of Analytical Chemistry*, **2020**.

24. Trineeva, O. V., Safonova, E. F., Sinkevich, A. V. and Slivkin, A. I. (2015) Assay of amino acids in medicinal plants by TLC (using stinging nettle leaves and common sea buckthorn fruits as examples). *Pharmaceutical Chemistry Journal*, **49(5)**, 323–328.

25. Rebiai, A., Seghir, B. B., Hemmami, H., Zeghoud, S., Siham, T., Kouadri, I. and Brahmia, F. (2021) Clustering and Discernment of bee pollen using an image analysis system. *Algerian Journal of Chemical Engineering AJCE*, **1(2)**, 41–48.

26. Rebiai, A., Hemmami, H., Zeghoud, S., Ben Seghir, B., Kouadri, I., Eddine, L. S. and Abbas, N. (2022) Current Application of Chemometrics Analysis in Authentication of Natural Products: A Review. *Combinatorial Chemistry & High Throughput Screening*, **25(6)**, 945–972.

27. Zeghoud, S., Hemmami, H., Rebiai, A. and Ben Seghir, B. (2022) Using FTIR spectroscopy and chemometrics for classifying of algerian medicinal plant species. *Vegetos*, **35(2)**, 298–305.

28. Rebiai, A., Ben Seghir, B., Hemmami, H., Zeghoud, S., Belfar, M. L. and Kouadri, I. (2021) Determination of some phenolic acids in Algerian propolis. *Ovidius University Annals of Chemistry*, **32**, 120–124.

29. Hemmami, H., Ben Seghir, B., Ben Ali, M., Rebiai, A., Zeghoud, S. and Brahmia, F. (2020) Phenolic profile and antioxidant activity of bee pollen extracts from different regions of Algeria. *Ovidius University Annals of Chemistry*, **31**, 93–98.

30. Zeghoud, S., Rebiai, A., Hemmami, H., Ben Seghir, B., Elboughdiri, N., Ghareba, S. and Abbas, N. (2021) ATR–FTIR spectroscopy, HPLC chromatography, and multivariate analysis for controlling bee pollen quality in some Algerian regions. *ACS omega*, **6(7)**, 4878-4887.

31. Wang, L. and Mizaikoff, B. (2008) Application of multivariate data-analysis techniques to biomedical diagnostics based on mid-infrared spectroscopy. *Analytical and bioanalytical chemistry*, **391(5)**, 1641–1654.

32. Alaerts, G., Pieters, S., Logie, H., Van Erps, J., Merino-Arévalo, M., Dejaegher, B. and Vander Heyden, Y. (2014) Exploration and classification of chromatographic fingerprints as additional tool for identification and quality control of several Artemisia species. *Journal of Pharmaceutical and biomedical Analysis*, **95**, 34–46.

33. Tistaert, C., Dejaegher, B., Hoai, N. N., Chataigné, G., Rivière, C., Hong, V. N. T. and Vander Heyden, Y. (2009) Potential antioxidant compounds in Mallotus species fingerprints. Part I: Indication, using linear multivariate calibration techniques. *Analytica chimica acta*, **649(1)**, 24–32.

34. Yi, W. S., Cui, D. S., Li, Z., Wu, L. L., Shen, A. G. and Hu, J. M. (2013) Gastric cancer differentiation using Fourier transform near-infrared spectroscopy with unsupervised pattern recognition. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, **101**, 127–131.

35. Granato, D., Santos, J. S., Escher, G. B., Ferreira, B. L. and Maggio, R. M. (2018) Use of principal component analysis (PCA) and hierarchical cluster analysis (HCA) for multivariate association between bioactive compounds and functional properties in foods: A critical perspective. *Trends in Food Science & Technology*, **72**, 83–90.

36. Orzel, J., Daszykowski, M., Kazura, M., de Beer, D., Joubert, E., Schulze, A. E. and Walczak, B. (2014) Modeling of the total antioxidant capacity of rooibos (Aspalathus linearis) tea infusions from chromatographic fingerprints and identification of potential antioxidant markers. *Journal of Chromatography A*, **1366**, 101–109.

37. Awuchi, C. G., Twinomuhwezi, H. and Awuchi, C. G. (2022) Hyphenated techniques in *Analytical Techniques in Biosciences, ed: Elsevier*, 125–145.

38. Sarvarian, M., Jafarpour, A., Awuchi, C. G., Adeleye, A. O. and Okpala, C. O. R. (2022) Changes in Physicochemical, Free Radical Activity, Total Phenolic and Sensory Properties of Orange (Citrus sinensis L.) Juice Fortified with Different Oleaster (Elaeagnus angustifolia L.) Extracts. *Molecules*, **27(5)**, 1530.

39. Awuchi, C. G. and Twinomuhwezi, H. (2021). The medical, pharmaceutical, and nutritional biochemistry and uses of some common medicinal plants. *Medicinal and Aromatic Plants of the World; Encyclopedia of Life Support Systems (EOLSS)*.

40. Natumanya, P., Twinomuhwezi, H., Igwe, V. S., Maryam, S. and Awuchi, C. G. (2021) Effects of Drying Techniques on Nutrient Retention and Phytochemicals in Selected Vegetables. *European Journal of Agriculture and Food Sciences*, **3(2)**, 5–14.

Abbreviations :

PCA	Principal Component Analysis
HCA	Hierarchical Clustering Analysis
PLS	Partial Least Squares method
PLS-DA	Partial Least Squares Discriminant Analysis
RMSECV	Root Means Square Error of Cross-Validation
RMSEE	Root Mean Squared Error of Estimation