



DETERMINATION OF EXTRACTIVE VALUES BY UV/VIS SPECTROPHOTOMETRY COUPLED WITH CHEMOMETRIC METHODS: CASE STUDIES OF TURMERIC, ANDROGRAPHIS AND ROSELLE

Tasamaporn Sukwattanasinit¹, Prapai Wongsinkongman², Panadda Phattanawasin¹,
Jankana Burana-Osot¹, Worakanya Narakornwit¹, Uthai Sotanaphun^{1,*}

¹ Department of Industrial Pharmacy, Faculty of Pharmacy, Silpakorn University, Sanamchandra Palace Campus, Nakhon Pathom

² Department of Medical Sciences, Ministry of Public Health, Nonthaburi

* Corresponding author: sotanaphun_u@su.ac.th

ABSTRACT

An extractive value is a quality control parameter of herbal raw materials. Its determination process requires patience and consumes a lot of electrical energy. This study explored the correlation between the extractive values and UV/Vis spectral data. The spectroscopic data was analyzed using partial-least square regression and multiple linear regression to estimate the extractive values. Turmeric, andrographis, and roselle were selected as case studies due to the diverse polarities exhibited by their chemical components. The predictive models for the ethanol-soluble extractives of all herbs, as well as the water-soluble extractives of andrographis and roselle, were effectively developed. Nevertheless, the model for the water extractive of turmeric did not yield successful results. The effectiveness of the chemometric approach relied on the chemical composition of the individual herbs. The reliability of the developed models was assessed by the acceptable root mean square error of cross-validation and root mean square error of prediction. The accuracy of each model was greater than 94%. This study presented an innovative concept that had the potential applicability to other herbs.

Keywords: extractive value, chemometric, UV/Vis spectrophotometry

Received: 25 March 2024; Revised: 8 May 2024; Accepted: 9 May 2024

Introduction

The quality of herbal raw materials is the initial important factor in the quality of herbal products. The pharmaceutical sector primarily follows the quality regulations addressed in the pharmacopoeia of each country. In Thailand, the Thai Herbal Pharmacopoeia (THP) was established in 1989,¹ and since then, an increasing number of herbal monographs that adhere to the generally accepted WHO guidelines² have been gradually released. However, many Thai herbal monographs still lack the assay method for quantifying active constituents due to the difficulty in accurately identifying the bioactive compounds. In such cases, the quantity of extractable matter is employed as a general test criterion,² with THP¹ providing standardized values of extractable matter in all monographs.

The quantity of extractable matter, also known as the extractive value, refers to the total weight of substances that can be extracted using a specified solvent, typically water or ethanol. This parameter has received significant attention in research publications.³⁻⁷ Determining the extractive value is a straightforward method that only requires minimal chemicals but takes time and patience. Furthermore, achieving a constant weight by drying the herbal extract demands an extensive amount of electrical energy.

Therefore, this study introduces a novel application of the chemometric method for determining extractive values using UV/Vis spectral data for the first time. The principle of this method involved establishing a spectrum of the extract prepared from a suitable solvent, analyzing the wavelengths relevant to the extractive value, and constructing a regression using chemometric methods to calculate the extractive value of unknown samples. Partial Least Square Regression (PLS) and Multiple Linear Regression (MLR) were the selected chemometric methods. PLS was the mathematic algorithm used to concurrently abstract the extractive value matrix (Y) and spectroscopic data matrix (X) into a reduced number of factors, represented by the

X-loading matric (P) and Y-loading matrix (Q), along with the score matrix (T). The mathematical formulas are expressed as $X = TP + E$ and $Y = TQ + F$, where E and F are residual matrices. MLR was employed to establish a relationship between the matrix of the extractive value (Y) and the matrix of spectroscopic data of the selected wavelengths (X) using a regression method. The mathematical equation is expressed as $Y = BX$, where B is coefficient matrix.⁸

Three herbs, turmeric (the dried rhizome of *Curcuma longa* L.), andrographis (the aerial part of *Andrographis paniculata* (Burm. f.) Nees), and roselle (the dried calyx and epicalyx collected during the fruiting of *Hibiscus sabdariffa* L.), were used as case studies due to the diverse range of polarities present in their chemical components. For spectral analysis, methanol extracts of turmeric and andrographis were utilized. Methanol was chosen for turmeric due to the predominance of low-polarity chemical compositions, while its use in andrographis was based on the diverse polarity of its chemical constituents. In contrast, a 50% methanol solution was employed for roselle, which contains highly polar chemical constituents. Consequently, it was anticipated that the acquired spectrum would reflect the absorbance of all chemical constituents within the herb, enabling the prediction of both ethanol-soluble and water-soluble extractive values through chemometric analysis. This approach helps conserve both human resources and electrical energy, providing an alternative method for in-house or in-process quality control.

Materials and Methods

Plant Materials

Forty-two samples of turmeric and thirty-one samples of andrographis were obtained from the Medicinal Plant Research Institute, Department of Medical Sciences, Ministry of Public Health, from 2008 to 2010. Thirty-one samples of roselle throughout Thailand were collected from 2003 to 2004. All voucher specimens are deposited at the

Herbarium of the Faculty of Pharmacy, Silpakorn University, Thailand. All samples were ground to powder, passed through a sieve mesh 0.250 mm and stored at 4°C. They were randomly divided into calibration, validation, and test sets as shown in Table 1. The extractive value and the UV/Vis spectrum of each sample were examined soon after the samples were received.

Determination of Extractive Value

Water-soluble and ethanol-soluble or 85% ethanol-soluble extractives of all samples were analyzed according to their THP monographs.¹ Five grams of the air-dried, powdered sample was accurately weighed and macerated with 100.0 ml of solvent in a closed flask for 24 hours. It was shaken frequently during the first 6 hours and then allowed to stand for 18 hours. The extract was filtered rapidly, and 20.0 mL of the filtrate was evaporated to dryness on a water-bath and further dried at 105°C in a hot-air oven to constant weight. Calculate the percentage of extractive with reference to the air-dried sample.¹

UV/Vis Spectroscopic Analysis

Turmeric powder (30 mg) was macerated with methanol (25 mL) at room temperature for 24 hours. One mL of the extract was diluted to 25 mL with methanol. Andrographis powder (80 mg) was macerated with methanol (25 mL) at room temperature for 24 hours. One mL of the extract was diluted to 10 mL with methanol. Roselle powder (1 g) was sonicated with 5 mL of 50% methanol for 15 min. The extract (0.2 mL) was diluted with 0.8 mL of 50% methanol and the solution (0.3 mL) was further diluted with 4.9 mL of 50% methanol. UV/Vis spectra in the range of 190–1100 nm of the final solutions of all samples were collected by a UV/Vis spectrophotometer (Agilent 8453 Model G 11C3A, USA) (Figure 1). All process was duplicated.

Data Analysis

This study was an analysis of measuring data (extractive value and UV/Vis spectrum) accumulated over several years earlier. PLS and MLR were carried

out using the Unscramble 9.8[®] (Camo Process AS, Norway). Before data analysis, spectral data were baseline corrected by subtraction with the absorbance at 900 nm. The efficiency of the model was validated through the error of prediction during full cross-validation of calibration and validation set. The results were presented as RMSECV (root mean square error of cross-validation) and RMSEP (root mean square error of prediction), respectively. These parameters can range from zero to positive infinity and use the same units as extractive value. A value of 0 means that the predicted values perfectly matched the measured values. Correlation (R) between measured and predicted values during validation also evaluated. The chosen model underwent further %accuracy assessment using the test set.

Results and Discussion

The extractive values of turmeric, andrographis and roselle samples were analyzed according to the monographs in the Thai Herbal Pharmacopoeia, with the results presented in Table 1. Figure 1 shows the UV/Vis spectra of 50% methanol extracts of roselle, as well as methanol extracts of turmeric and andrographis in the wavelength range showing light absorption. The relationship between their extractive values and spectral data was analyzed by PLS and MLR methods, and the models were established as follows.

Prediction Models of Turmeric

PLS models of ethanol-soluble extractive based on UV/Vis (model #1, 200–600 nm) and UV (model #2, 200–300 nm) of the methanol extract were developed. Both models provided comparable predictive power (Table 2). This indicated that information based on visible spectra was not necessary. The results obtained from the model using the visible data range (model #3, 300–600 nm) gave unsatisfactory outcomes. This could be attributed to the fact that curcuminoids and volatile oils are the major chemical constituents of turmeric⁹, and both can absorb UV light. However, only curcuminoids

responded to yellow visible light. Thus, despite curcuminoids showing a strong absorption band in the visible region, UV data alone proved adequate for the model. The selected PLS model #2 is shown in Figure 2. MLR models were further constructed from absorbance at λ_{max} . Similar to PLS, MLR including $A_{203,233,420}$ (model #4) gave comparable predictive efficiency to the model without A_{420} (model #5), and model #6 which used only A_{420} had a highly predictive error. Then A_{420} was unnecessary, resulting in the following simplified model.

ethanol-soluble extractive = $0.347696 + 13.922A_{203} + 14.320A_{233}$ (model #5).

The selected PLS model #2 and MLR model #5 were applied to the test set and gave satisfactory results (Table 3). On the contrary, a predictive model of water-soluble extractive was unsuccessfully developed. Most of the water-soluble substances of turmeric are polysaccharides that generally have no chromophore. Then it was impossible to correlate the relation between UV/Vis spectrum and water-soluble extractive.

Table 1 Number of samples (N) and extractive values (%w/w)* of each sample group.

Sample set	Turmeric			Andrographis			Roselle		
	N	Ethanol-soluble	Water-soluble	N	85% Ethanol-soluble	Water-soluble	N	Ethanol-soluble	Water-soluble
Calibration	23	6.8–40.0	11.5–22.4	16	13.1–29.0	14.4–26.6	17	4.2–19.8	41.7–55.6
Validation	12	9.8–24.9	12.8–20.3	8	17.5–26.4	21.4–27.1	8	6.0–17.2	42.0–52.1
Test	7	14.2–22.2	16.1–19.8	7	15.5–26.1	18.2–25.5	6	6.1–14.6	44.5–54.2

* Data presented as minimum to maximum values

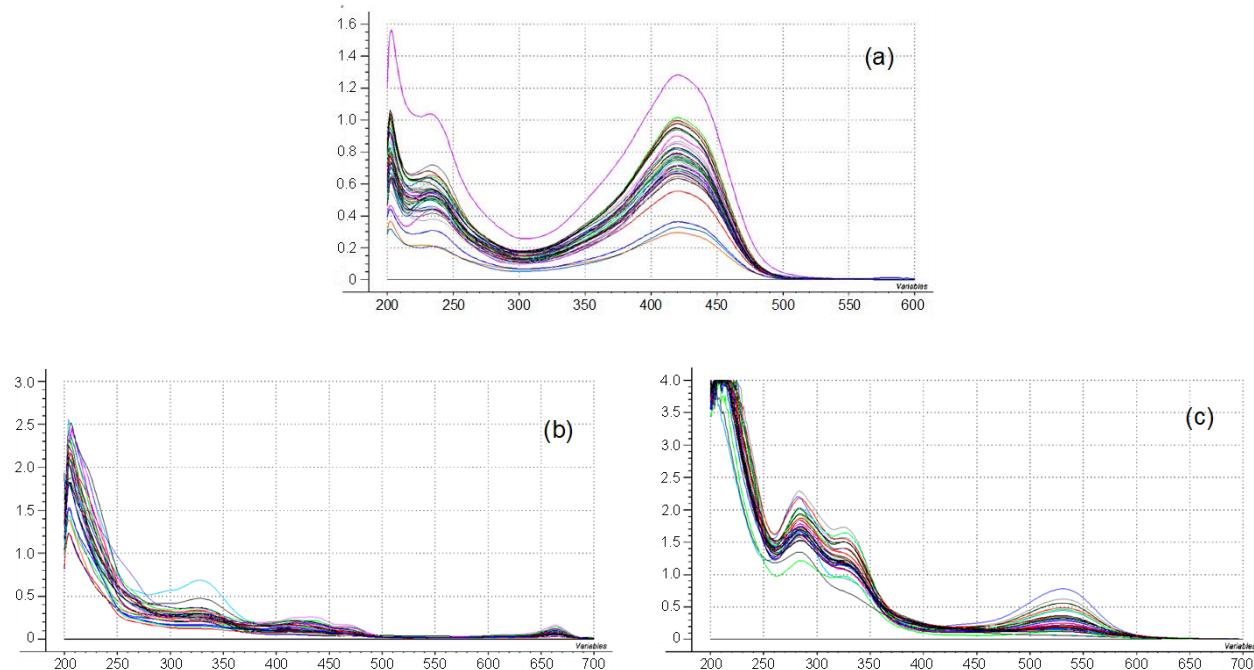


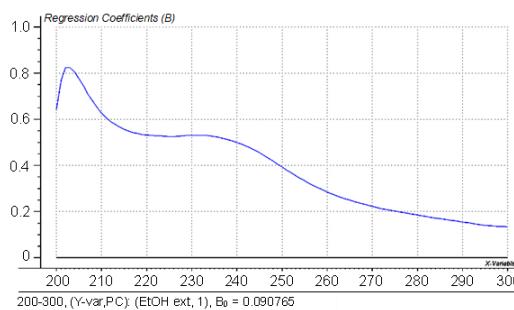
Figure 1 Ultraviolet/visible spectra of methanol extracts of all samples of (a) turmeric, (b) andrographis and (c) roselle.

Table 2 Validation parameters of the prediction models of ethanol-soluble extractive of turmeric.

Model No.	Method	Wavelength (nm)	Number of PLS factor	RMSECV	R (LOOCV) ^a	RMSEP	R (test) ^b
1	PLS	200–600	2	1.8952	0.9490	1.5836	0.9347
2	PLS	200–300	1	1.8106	0.9535	1.5838	0.9366
3	PLS	300–600	2	2.4378	0.9140	2.4373	0.8791
4	MLR	203,233,420	-	2.0454	0.9399	1.7298	0.9155
5	MLR	203,233	-	1.9238	0.9472	1.7102	0.9187
6	MLR	420	-	3.2830	0.8360	2.5579	0.8816

^aCorrelation between measured and predicted values during leave one out cross-validation

^bCorrelation between measured and predicted values of the validation set

**Figure 2** PLS prediction model of ethanol-soluble extractive of turmeric (model #2).**Table 3** Ethanol-soluble extractive (%w/w) of tested samples of turmeric using the prediction models.

Tested sample	Ethanol-soluble extractive (%w/w)		
	Measured value	PLS (200–300 nm)	MLR (203,233 nm)
1T	22.22	20.69	19.88
2T	21.60	19.23	18.88
3T	14.17	15.04	15.00
4T	17.03	18.30	18.41
5T	15.46	16.12	16.05
6T	15.14	13.69	13.10
7T	22.10	23.28	23.45
	% Accuracy	99.40±8.13	98.18±9.83

Prediction Models of Andrographis

PLS models based on UV/Vis data (200–700 nm) and UV data (200–400 nm) of methanol extract were developed to predict 85% ethanol-soluble

extractive (model #7 and #8, respectively), and water-soluble extractive (model #11 and #12, respectively). The results (Table 4) were in the same manner as turmeric. The models including visible spectral data

did not show improved validation results compared to those without it. Major chemical constituents of andrographis were diterpene lactones (andrographolide derivatives), flavonoids, and other phenolic compounds.¹⁰ These compounds are colorless and are capable of absorbing UV light. The minor absorption bands observed in the visible region are attributed to plant pigments,¹¹ which are less soluble in water and 85% ethanol. Hence, data from the visible region was not important. Models #8 and #12 were selected to predict ethanol-soluble and water-soluble extractives, respectively (Figure 3). Surprisingly, the pattern of both models was slightly similar, but the coefficients of the water-soluble extractive model were much lower. This suggested that the UV-active compositions of both extractives were not much different, but they were less concentrated in water-soluble extractive. Additionally, water is known to primarily dissolve highly polar compounds, such as primary metabolites, which generally do not have UV-chromophores. Then having a very high intercept

value of water-extractive model #12 ($B_0 = 10.3762$) compared to 85% ethanol-extractive model #8 ($B_0 = 1.8661$) might be due to this reason. The models were simplified by MLR using only A_{\max} at 207, 330, and 415 nm (models #9-10 and 13-14). The predictive powers of these models were comparable to PLS and confirmed that A_{415} or visible data was not necessary. The models for 85% ethanol-soluble and water-soluble extractives were as follows:

$$85\% \text{ ethanol-soluble extractive} = 2.262318 + 7.844A_{207} + 9.722A_{330} \text{ (model #10),}$$

$$\text{water-soluble extractive} = 10.020055 + 6.234A_{207} + 0.130A_{330} \text{ (model #14).}$$

Similarly to PLS, when compared with the model for 85% ethanol-soluble extractive, the intercept value of the model for water-soluble extractive was much higher, while its coefficient values were smaller. All selected PLS models #8, #12 and MLR models #10, #14 were applied to the test set and gave satisfactory results (Table 5).

Table 4 Validation parameters of the prediction models of 85% ethanol-soluble extractive and water-soluble extractive of andrographis.

Model No.	Method	WL	Number of PLS factor	RMSECV	R (LOOCV)	RMSEP	R (test)
85% Ethanol-soluble extractive							
7	PLS	200–700	1	2.5581	0.8252	1.2011	0.9690
8	PLS	200–400	1	2.5395	0.8280	1.2330	0.9679
9	MLR	207,330,415	-	3.2558	0.7056	1.0787	0.9583
10	MLR	207,330	-	2.7787	0.7897	1.0474	0.9596
Water-soluble extractive							
11	PLS	200–700	1	2.3229	0.6811	1.1338	0.9053
12	PLS	200–400	1	2.3208	0.6820	1.1342	0.9068
13	MLR	207,330,415	-	2.7876	0.5553	1.0389	0.8987
14	MLR	207,330	-	2.5954	0.6227	1.0576	0.8911

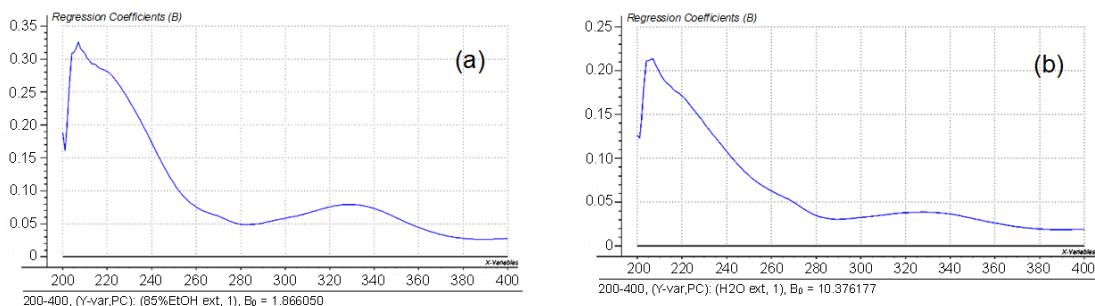


Figure 3 PLS models of (a) 85% ethanol-soluble extractive (%w/w) and (b) water-soluble extractive (%w/w) of tested samples of andrographis.

Table 5 85% Ethanol-soluble extractive (%w/w) and water-soluble extractive (%w/w) of tested samples of andrographis using the prediction models.

Tested sample	85% Ethanol-soluble extractive (%w/w)			Water-soluble extractive (%w/w)		
	Measured value	PLS (200–400 nm)	MLR (207,330 nm)	Measured value	PLS (200–400 nm)	MLR (207,330 nm)
1A	17.64	15.78	16.40	21.37	19.24	19.07
2A	26.36	24.26	24.17	25.02	24.65	24.70
3A	25.2	25.79	24.06	25.45	25.67	24.61
4A	23.19	25.43	24.03	21.7	25.34	25.31
5A	17.48	18.43	18.83	22.43	20.95	20.90
6A	21.19	20.80	20.20	25.92	22.43	22.02
7A	25.12	21.27	21.93	27.11	22.80	23.69
% Accuracy	97.39 ± 9.07	96.30 ± 7.06		95.75 ± 11.07	95.26 ± 10.64	

Prediction Models of Roselle

Due to the predominance of polar compounds in roselle, such as anthocyanins, phenolic compounds, plant acids and mucilage,¹² 50% methanol solution was prepared for UV/Vis spectrum to ensure their solubility. In contrast to turmeric and andrographis, both UV and visible spectral information (model #15, 250–650 nm) were crucial for the prediction of ethanol-soluble extractive by the PLS model (Table 6). The models constructed with UV data (250–425 nm, model #16) or visible data (425–650 nm, model #17) alone, provided poor efficiency. PLS model #15 is shown in Figure 4(a). This model contained numerous spectral data and required up to 9 PLS factors to optimize the validation result.

Then the model was very complicated and it possibly caused misleading predictive results of unknown samples.¹³ Nevertheless, upon application to the test set, it yielded slightly satisfactory outcomes (Table 7). To simplify the model, MLR was attempted to develop from λ_{max} at 280, 325, and 530 nm. Moreover, λ_{380} which gave a high coefficient value in the PLS model, was also included. However, neither model provided an acceptable validation result (Data did not show). As discussed above, these limited number of wavelengths were insufficient to explain the complex relationships between the UV/Vis spectrum and the ethanol-soluble extractive of roselle.

PLS models of water-soluble extractives based on UV/Vis, UV, and visible data (models #18, #19, and #20, respectively) were developed. Their validation results were quite similar. However, the validation result of model #19, based on only UV data, was slightly poor. Considering model #18 (Figure 4(b)), the coefficient values in the visible range were positive, whereas most UV data were negative. The visible data was the absorption band of reddish anthocyanin. Anthocyanin was highly soluble in water and thus responded mainly to water-soluble extractive. In the UV region, besides anthocyanin, there were also other UV-active compounds. The composition of these UV-active compounds of 50%methanol extract and water-soluble extract might have been considerably different. Therefore, it was inappropriate to include UV data in the model. PLS model #20 using only visible data was selected (Figure 4(c)). MLR also confirmed the inefficiency of

UV data (model #22, $A_{280,325}$). The predictive powers of model #21 ($A_{280,325,530}$) and model #23 (A_{530}) were comparable. As discussed above and to simplify the model, model #23 was chosen. Its model was as follows:

$$\text{water-soluble extractive} = 44.748875 + 10.644A_{530} \text{ (model #23).}$$

A point of concern was the high intercept values ($B_0 = 55.2200$ and 44.7489) of both selected models #20 and #23, respectively. These constants corresponded to the substantial amount of water-soluble polysaccharides, e.g., mucilage and pectin, which were found up to 10–30% in roselle¹¹. Polysaccharides do not absorb light and could therefore interfere with the prediction. As a result, the validated results for both models were slightly poor, and upon application to the test set, they provided only slightly satisfactory results (Table 7).

Table 6 Validation parameters of the prediction models of ethanol-soluble extractive and water-soluble extractive of roselle.

Model No.	Method	Wavelength (nm)	Number of PLS factor	RMSECV	R (LOOCV)	RMSEP	R (test)
Ethanol-soluble extractive							
15	PLS	250–650	9	2.2429	0.8527	1.9519	0.9111
16	PLS	250–425	6	3.2271	0.6453	4.2114	0.2677
17	PLS	425–650	7	2.5063	0.7545	3.1250	0.5732
Water-soluble extractive							
18	PLS	250–650	2	3.3709	0.3592	2.6715	0.5972
19	PLS	250–425	2	3.2370	0.4015	3.4558	0.2815
20	PLS	425–650	2	3.1153	0.4615	2.6974	0.5438
21	MLR	280,325,530	-	3.9515	0.1650	3.1167	0.5218
22	MLR	280,325	-	3.8913	-0.2645	3.4313	-0.5797
23	MLR	530	-	3.3033	0.3226	3.1871	0.3603

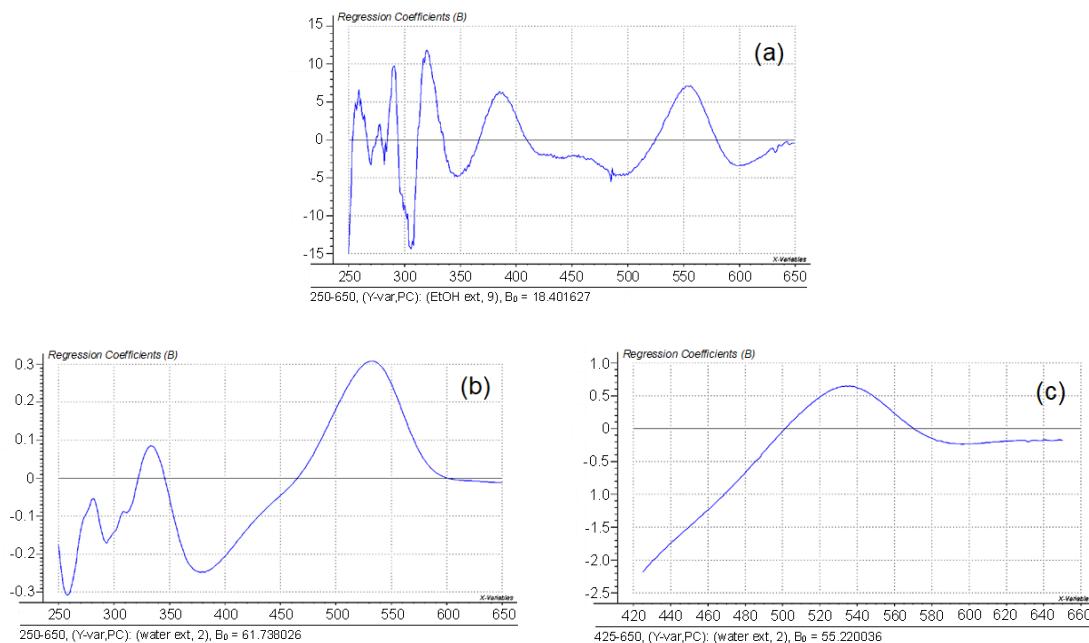


Figure 4 PLS models of (a) ethanol-soluble extractive (model #15), and (b) (c) water-soluble extractive (models #18 and 20) of roselle.

Table 7 Ethanol-soluble extractive (%w/w) and water-soluble extractive (%w/w) of tested samples of roselle using the prediction models.

Tested sample	Ethanol-soluble extractive (%w/w)		Water-soluble extractive (%w/w)		
	Measured value	PLS	Measured value	PLS	MLR
		(250–650 nm)		(425–650 nm)	(530 nm)
1R	14.64	12.58	50.28	50.10	47.84
2R	11.45	11.16	50.59	50.56	47.40
3R	11.75	11.10	54.22	47.75	48.09
4R	6.90	7.36	46.20	45.09	46.56
5R	6.13	5.52	44.54	44.70	46.80
6R	6.64	6.15	46.81	45.37	46.62
%Accuracy		94.53 ± 7.12		97.09 ± 4.63	97.16 ± 5.82

Conclusion

This study proposed the use of the chemometric methods (PLS and MLR) to determine both ethanol-soluble and water-soluble extractives using only a single UV/Vis spectrum. Models for both extractive parameters of andrographis and roselle were successfully established, while only the ethanol-soluble extractive model of turmeric was satisfactorily

developed. All models demonstrated an accuracy of more than 94%. The PLS and MLR models provided comparable efficiency and could be chosen for future use according to convenience. The relationship between extractive values and UV/Vis spectral data depended on the chemical compositions of individual herbs and the light-absorbing properties of the soluble constituents in the extracts. The

concept of this approach could be applied to analyze the extractive values of other herbs. It was straightforward, rapid, energy-saving, and could be utilized to assess the quality of herbal raw materials in routine tasks.

Acknowledgments

The authors would like to thank Faculty of Pharmacy, Silpakorn University, for the provision of laboratory facilities, and the Medicinal Plant Research Institute, Department of Medical Sciences, Ministry of Public Health, for samples of turmeric and andrographis, and data of their extractive values.

Conflict of Interest

The authors declare that there is no conflict of interest.

References

1. Thai Pharmacopoeia Committee. Thai Herbal Pharmacopoeia [Internet]. 2024 [cited 2024 March 15]. Available from: <https://bdn.go.th/thp/home>
2. World Health Organization. Quality control methods for medicinal plant materials [Internet]. 1998 [cited 2024 March 15]. Available from: https://www.who.int/docs/default-source/medicines/norms-and-standards/guidelines/quality-control/quality-control-methods-for-medicinal-plant-materials.pdf?sfvrsn=b451e7c6_0
3. Sharma R, Amin H, Shukla VJ, Kartar D, Galib R, Prajapati PK. Quality control evaluation of *Guduchi Satva* (solid aqueous extract of *Tinospora cordifolia* (Willd.) Miers): An herbal formulation. *Int J Green Pharm.* 2013;7(3):258–63.
4. Hartanti D, Theeravit J. Quality control of crude drugs and capsules of turmeric (*Curcuma longa* L.). *Pharm J Indones.* 2018;15(2):257–66.
5. Pannu A, Goyal RK. Pharmacognostical characterization of whole *Scutellaria barbata* herb for quality control assessment. *Innov Int J Med Pharm Sci.* 2021; 6(4):1–5.
6. Zambare AV, Chakraborty GS, Banerjee SK. Pharmacognostic studies of potential herb - *Tridax procumbens* Linn. *IJPSR.* 2010;1(9):58–62.
7. Chandel HS, Pathak AK, Tailang M. Standardization of some herbal antidiabetic drugs in polyherbal formulation. *Pharmacognosy Res.* 2011;3(1):49–56.
8. Brereton RG. Chemometric: Data analysis for the laboratory and chemical plant. Chichester: John Wiley & Sons; 2003.
9. Fuloria S, Mehta J, Chandel A, Sekar M, Rani NNIM, Begum MY, et al. A comprehensive review on the therapeutic potential of *Curcuma longa* Linn. in relation to its major active constituent curcumin. *Front Pharmacol.* 2022;13:820806.
10. Subramanian R, Asmawi MZ, Sadikun A. A bitter plant with a sweet future? A comprehensive review of an oriental medicinal plant: *Andrographis paniculata*. *Phytochem Rev.* 2012;11:39–75.
11. Ashenafi EL, Nyman MC, Shelley JT, Mattson NS. Spectral properties and stability of selected carotenoid and chlorophyll compounds in different solvent systems. *Food Chem Adv.* 2023;2:100178.
12. Da-Costa-Rocha I, Bonnlaender B, Sievers H, Pischel I, Heinrich M. *Hibiscus sabdariffa* L. – A phytochemical and pharmacological review. *Food Chem.* 2014;165:424–43.
13. Tanaka K, Kuba Y, Sasaki T, Hiwatashi F, Komatsu K. Quantitation of curcuminoids in curcuma rhizome by near-infrared spectroscopic analysis. *J Agric Food Chem.* 2008;56(19):8787–92.