

Development and validation of global prediction models for monitoring the manufacturing process of herbal medicine by ultraviolet spectroscopy

Jie Zhao^{1,2}, Zimei Zhou^{1,2}, Fang Zhao^{1,2}, Xu Yan^{1,2}, Jianyang Pan^{1,2}, Haibin Qu^{1,2,*}

Abstract

Objective: Process monitoring for traditional Chinese medicine (TCM) preparations is necessary to ensure the quality of the product. A typical pharmaceutical process of TCM preparations consists of multiple manufacturing units, such as ethanol precipitation, concentration, and water precipitation, among others. Compared with the traditional practice of one prediction model for one unit, the global model covers the variation from samples with different backgrounds or processes and can be used to monitor intermediates from substeps.

Methods: We used ultraviolet (UV) spectroscopy to establish global models for a typical TCM preparation—Danhong injection. The concentrations of danshensu, protocatechualdehyde, rosmarinic acid, salvianolic acid A, salvianolic acid B, and hydroxyl safflor yellow A and the total phenolic and total sugar contents were quantified for every intermediate from operation units of Danhong injection. New samples prepared by mixing different intermediates were introduced for the calibration set to cover more variations. An accuracy profile was employed to validate the developed method from the aspects of specificity, trueness, precision, accuracy, linearity, and robustness.

Results: The developed models showed a high determination coefficient (R^2) value up to 0.97 and a low root-mean-square error of the prediction set. Five components of the models passed all validation tests, whereas the total sugar was not suitable for modeling with UV and did not apply to the whole process.

Conclusions: This study indicates that the global models of UV spectroscopy for the quantitative determination of phenolic acids are feasible and reliable with a simple, rapid, and non-destructive method.

Graphical abstract: <http://links.lww.com/AHM/A22>.

Keywords: Accuracy profile, Danhong injection, Global model, Model validation, UV spectroscopy

Introduction

The manufacturing process of Chinese herbal preparations consists of a series of unit operations, such as extraction, concentration, ethanol precipitation, and water precipitation, among others^[1–2]. Each unit operation involves a set of variations in raw materials and process parameters. Therefore, the quality of the intermediate should be closely monitored to ensure consistent quality of the final products from batch to batch^[3], which matches the requirements of the International Conference on Harmonization (ICH) Q8 guidelines introduced by the United States Food and Drug Administration^[4].

Spectroscopy methods, such as ultraviolet (UV) spectroscopy, near-infrared (NIR) spectroscopy, and Raman spectroscopy, have been used as powerful process analytical technology tools for the fast and nondestructive quantification of multiconstituent mixtures and process monitoring^[5–6]. Because the majority of active ingredients in Chinese herbal medicines (CHMs) have strong UV absorption, UV spectroscopy is suitable for measuring active ingredients in Chinese herbal preparations^[7–8]. With the advantages of stability and inexpensiveness, UV spectroscopy is widely used for the quantitative analysis of phenolic acids, tannins, and flavones of CHMs off-line, at-line, or in-line^[9–10].

Chemometrics provides a very practical method for multivariable models in spectroscopic analysis^[11–12]. Generally, each unit operation in the pharmaceutical process requires an independent prediction model to monitor the variation. If any changes occur in the operating condition or material properties, a model update is required. However, even the preparation of a single CHM contains multiple processes, and thus, different models need to be established for different units, as the backgrounds between processes are relatively large, making model development a great challenge for UV spectroscopy.

The global model is a more adaptive model that covers variations in samples with different backgrounds or those produced with contrasting processes, and it increases the effectiveness and applicability of the model^[13–14]. Model validation needs to be carried out systematically according to ICH Q2 during its development^[15–16]. To evaluate the global model, the root-mean-square error of the

¹Pharmaceutical Informatics Institute, College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, China; ²Innovation Center in Zhejiang University, State Key Laboratory of Component-Based Chinese Medicine, Hangzhou, China

* Corresponding author. Haibin Qu, College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310058, China, E-mail: quhb@zju.edu.cn.

Copyright © 2022 Tianjin University of Traditional Chinese Medicine. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Acupuncture and Herbal Medicine (2022) 2:2

Received 16 August 2021 / Accepted 07 June 2022

<http://dx.doi.org/10.1097/HM9.0000000000000034>

prediction set (RMSEP), determination coefficient (R^2), residual prediction deviation (RPD), range error ratio, and other statistical tests^[17] are commonly used to confirm that there is no significant deviation between the measured values and the predicted values. As part of a proposal to harmonize validation protocols of quantitative analytical methods, the commission of the Société Française des Sciences et Techniques Pharmaceutiques (SFSTP) has introduced the accuracy profile as a decision tool in method validation to ensure the effectiveness of the method by assessing the risk of method failure^[18–22]. Although the accuracy profile was substantiated by NIR^[23–25] and UV spectroscopy^[26–27], to the best of our knowledge, there is currently no report on the validation of the global model with the accuracy profile.

Danhong injection, composed of *Salviae miltiorrhizae Radix et Rhizoma* (Danshen) and *Carthami Flos* (Honghua), has shown cardioprotective effects and is widely used for cardiovascular diseases^[28–29]. In this work, UV spectroscopy combined with a partial least squares (PLS) algorithm was used to develop global models containing a variety of intermediates—extraction, concentration, ethanol precipitation, dilution, and water precipitation—in the manufacturing process. The results showed that the models could quantify the concentrations of danshensu (DSS), protocatechualdehyde (PA), rosmarinic acid (RA), salvianolic acid A (SAA), salvianolic acid B (SAB), and hydroxyl safflor yellow A (HSYA) and total phenolic and total sugar contents for the Danhong injection. The accuracy profile was used for model validation from the aspects of specificity, trueness, precision, accuracy, linearity, and robustness using mixture samples independent from the calibration set. This study found that the different samples in multiple processes can be used for global modeling, and the specificity and robustness of the models were validated by the accuracy profile method.

Materials and methods

Samples for model development

Salviae miltiorrhizae Radix et Rhizoma (Danshen), *Carthami Flos* (Honghua), and original concentrated solutions were obtained from Shandong Buchang Pharmaceuticals (Shandong, China). The standard substances of DSS, PA, RA, SAA, and SAB were obtained from Shanghai Winherb Medical Technology Co., Ltd. HSYA was obtained from Chengdu Must Bio-Technology Co., Ltd. Acetonitrile and formic acid were high-performance liquid chromatography (HPLC)-grade and supplied by Merck (Darmstadt, Germany). Ultrapure water was produced by a Milli-Q water purification system (Miford, MA, USA) in the laboratory. An Agilent 1100 HPLC system equipped with Chemstation Software (Agilent Technologies, Santa Clara, USA) was used for quantitative analysis. Spectral scans and other absorbance determinations were conducted on an Agilent Cary 60 UV-Vis spectrophotometer (Agilent Technologies, Santa Clara, USA).

Sample preparation

Samples for model development

The intermediates of each process were annotated A, B, C, D, E, and F. To ensure that the samples collected from the

extraction process (A) had sufficient variation, a three-factor Box-Behnken design (BBD) experiment covering 18 runs was carried out^[30]. Seventeen samples were obtained from process A, whereas sample No. 1 was discarded owing to improper preservation. The experimental design is shown in Table 1.

The ethanol precipitation process (B) consisted of two steps. First, the original concentrated solution was divided into 10 parts with the same volume. Then, ethanol was added into the concentrated solution at different amounts *via* a peristaltic pump at 50 rpm, and the mixture was continuously stirred. It was then kept at 4°C for 24 h. After filtration, 10 parts of alcohol precipitation solution (B₁) were collected. Then, one copy of B₁ was randomly removed and concentrated at 60°C under reduced pressure. The concentrated solution was divided into three parts. Different amounts of ethanol were added to the mixture and then centrifuged at 50 rpm followed by continuous stirring to achieve solutions with 80%, 82.5%, and 85% of the apparent alcohol concentration. The mixture was kept at 4°C for 24 h, and then, three parts of the alcohol precipitation (B₂) were collected.

Water precipitation samples (C) were made using B₂, similar to ethanol precipitation. B₂ was concentrated at 60°C under reduced pressure, and each part was equally divided into three parts. Then, nine samples of the first water precipitation solution (C₁) were collected. The three samples of C₁ were concentrated to 130 mL, 150 mL, and 163 mL under reduced pressure at 60°C and kept at 4°C for 24 h. After filtration, nine parts of secondary water precipitation solution (C₂) were collected. In summary, 18 water precipitation solutions were obtained.

The dilution (D) and concentration process (E) were conducted using samples from A. Three samples (No. 16, 17, 18) were arranged based on the BBD and diluted to half of the original concentration, and then, three diluent samples were prepared. To prepare sample E, three samples were randomly selected from A and mixed into

Table 1

Box-Behnken design for the extraction process.

No.	Extraction temperature (°C)	Boiling time (min)	Solvent times
1*	120	30	8
2	165	30	8
3	120	90	8
4	165	90	8
5	120	60	6
6	165	60	6
7	120	60	10
8	165	60	10
9	150	30	6
10	150	90	6
11	150	30	10
12	150	90	10
13	150	60	8
14	150	60	8
15	150	60	8
16	150	60	8
17	150	60	8
18	150	60	8

* No. 1 was discarded due to improper preservation.

three parts. The samples were concentrated at 60°C and diluted 4-, 7-, 10-, and 12-fold with the extraction solution. Then, 15 samples of the concentrated solution were collected.

To ensure that the global model covered as many variations as possible, another sample set F was prepared by mixing A, B, and C randomly, and eight samples of the mixture were obtained. Overall, we prepared 74 samples that contained all process intermediates of the Danhong injection preparation. The details of the sampling flow are shown in Figure 1, and the corresponding reference values are listed in Table 2.

Samples for model validation

An external validation set was constructed to evaluate the performance of the models. Because the PA concentration was relatively lower prior to the water precipitation process, we mixed samples C₁ and C₂ to validate the PA model, whereas other components were mixed using a fresh extract and sample C₁. The original mixture was diluted to six concentration levels. Each external validation set at individual levels was performed in triplicate for each series of experiments.

Reference method

HPLC analysis

The concentrations of DSS, PA, RA, SAA, SAB, and HSYA in the samples were determined by HPLC. Chromatography was achieved on a Waters Cortecs C₁₈ analytical column (100 mm × 4.6 mm, 2.7 μm) with a mobile phase consisting of (A) 0.4% HCOOH-H₂O and (B) 0.4% HCOOH-CH₃CN at a flow rate of 0.8 mL·min⁻¹. The gradient elution program was as follows: 0–9 min, 2% B; 9–27 min, 2–9% B; 27–40 min, 9–16% B; 40–53 min, 16–19% B; 53–63 min, 19–23% B; 63–65 min, 23–90% B. The UV detection wavelength was programmed as follows: 0–16 min, 280 nm; 16–32 min, 400 nm; 32–69 min, 280 nm. The column temperature was maintained at 40°C. The analysis method was verified by methodology, and the results were considered acceptable.

Total phenolic content determination

The total phenolic content in each sample was estimated using Folin-Ciocalteu method^[31]. Gallic acid was used as a standard in this work.

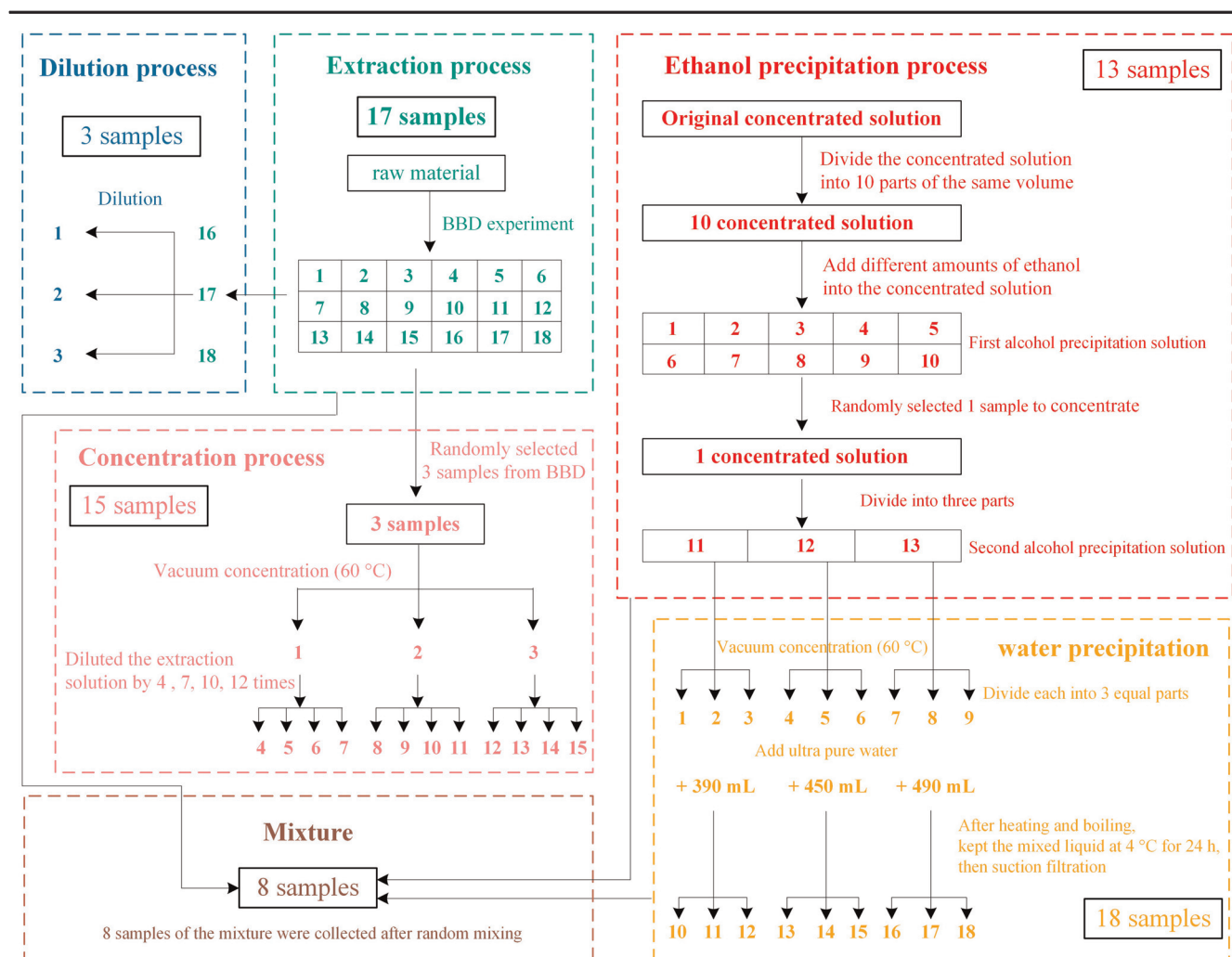


Figure 1. Sampling flow chart.

Table 2
Detailed information for the samples and reference values (after diluted).

Analyte	Sample numbers					
	A	B ₁ +B ₂	C ₁ +C ₂	D	E	F
	17	13	18	3	15	8
DSS (μg/g)			3.13–19.90			
PA (μg/g)			0.25–10.26			
RA (μg/g)			2.47–8.40			
SAB (μg/g)			19.18–124.67			
SAA (μg/g)			0–30.90			
HSYA (μg/g)			0.04–30.88			
Total phenolic (mg/g)			0.04–0.27			
Total sugar (mg/g)			0.09–1.37			

DSS: Danshensu; HSYA: Hydroxyl safflor yellow A; PA: Protocatechualdehyde; RA: Rosmarinic acid; SAA: salivianolic acid A; SAB: salivianolic acid B.

Total sugar content determination

Following the basic protocol^[32] of a modification method (ie, the phenol-sulfuric acid method), the diluted sample (1 mL) and 6% phenol solution (1 mL) were added to test the tube with a plug. Then, 5 mL of concentrated sulfuric acid was slowly added down the side of the tube. The tubes were then closed, vortexed, and incubated for 20 min at 100°C using a water bath. All tubes were allowed to cool to room temperature before reading the absorbance at 490 nm with ultrapure water replacing the sample as a blank and glucose as a standard.

Spectral acquisition

The spectra of samples were acquired in transmission mode with a 1 mm quartz cuvette using a UV-Vis spectrophotometer. The scanning range was 190 to 600 nm, with a resolution of 1 nm. Before scanning, the samples were diluted to an appropriate concentration to prevent overload in the spectrophotometer.

Model development

Fifty of the 74 samples were divided into the calibration set, and the remaining 24 were divided into the validation set with the Kennard-Stone method^[33]. Each set contained samples obtained from every intermediate step. The PLS regression method was used for global model development. The root-mean-square error of cross-validation (RMSECV) was used to determine the optimum number of principal components, and RMSEP was used for the selection of the data preprocessing method *via* the trial-and-error method. In addition, the R^2 of the validation set was adopted to evaluate the prediction performance of the models.

Accuracy profile

Validation of the developed method was performed using the total error approach as per SFSTP harmonization guidelines based on the accuracy profile and the β -expectation tolerance intervals (β -ETI). The β -ETI parameter is the interval within the acceptance limits λ where a proportion β of future measurements is expected

to be^[34]. If the β -ETI is included inside the range, the developed method is validated. The β -ETI is computed as follows:

$$\left[\hat{\mu}_j - Q_t\left(v; \frac{1+\beta}{2}\right) \sqrt{1 + \frac{1}{pn\beta_j^2} \hat{\sigma}_{IP,j}^2}, \hat{\mu}_j + Q_t\left(v; \frac{1+\beta}{2}\right) \sqrt{1 + \frac{1}{pn\beta_j^2} \hat{\sigma}_{IP,j}^2} \right] \quad (1)$$

$$v = \frac{(1+R)^2}{\frac{(R+1/n)^2}{p-1} + \frac{1-1/n}{pn}} \quad (2)$$

$$B_j = \sqrt{\frac{R_j + 1}{nR_j + 1}} \quad (3)$$

$$R_j = \frac{\hat{\sigma}_{B,j}}{\hat{\sigma}_{W,j}} \quad (4)$$

$$\hat{\sigma}_{IP,j}^2 = \hat{\sigma}_{W,j}^2 + \hat{\sigma}_{B,j}^2 \quad (5)$$

where $\hat{\mu}_j$ is the mean results of the concentration and $Q_t(v; \frac{1+\beta}{2})$ is the β quantile of Student's t distribution with v degrees of freedom. Moreover, p is the number of series, and n is the number of independent replicates per series.

Model validation

According to the method's validation protocol listed in ICH Q2, the developed method was examined from the aspects of specificity, trueness, precision, accuracy, linearity, and robustness. Since there is no specific requirement for spectroscopy method validation, specificity was determined according to the reference^[35].

Software

Sample division was performed using BWIQ 2.1.9 (B&W Tech, Inc., Newark, DE, USA). The regression model and data preprocessing were conducted on a SIMCA 14.1 (MKS Umetrics AB, Sweden). Accuracy profiles and other statistical treatments of method validation data were carried out using Microsoft Excel 2017 (Microsoft Inc, Redmond, USA) and Minitab 16 (Minitab Inc, Pennsylvania, USA).

Results and discussion

Modeling

Overview of the spectra

UV spectra of all 74 samples are displayed in Figure 2, and the spectral shapes of the samples were similar. The corresponding absorbance in the wavelength range of 190 to 215 nm was higher than 2, as many compounds exhibited terminal absorption. The analysis of UV spectra needs to be combined with chemometrics.

Global model construction

The PLS algorithm was used to develop global models of DSS, PA, RA, SAA, SAB, HSYA, total phenolic content, and total sugar content of Danhong injection intermedi-

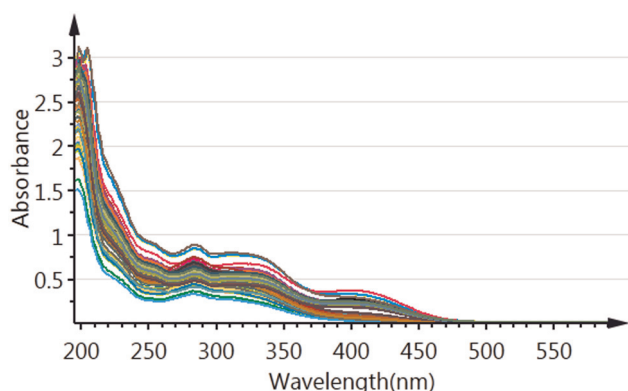


Figure 2. UV transmission raw spectra (190–600 nm) of 74 samples. UV: Ultraviolet.

ates. The preprocessing method for UV spectra was optimized by trial-and-error to obtain a sufficiently low RMSEP. Derivative and Savitzky–Golay filters were selected in this study to correct the baseline offsets and filter noise.

In the preliminary step to building the quantitative model of total phenolic and total sugar, the results displayed poor predictive performance. It could be inferred that the model might not apply to samples obtained from all intermediate operation processes. After sample screening and exploration, the total phenolic model was reconstructed without water precipitation samples ($C_1 + C_2$), as was done for the total sugar model without secondary water precipitation samples (C_2).

The results of the global models are shown in Table 3, including the RMSECV, RMSEP, and R^2 for all components. The R^2 values for both the calibration and validation sets were greater than 0.96, which showed a satisfactory correlation between the reference and the model predictions. The approximate value between RMSECV and RMSEP also indicated that no overfitting occurred. The correlation plots between the predicted value and the reference value are illustrated in Figure 3. The predicted concentrations were considerably close to the reference, which displayed a high correlation between them. The results showed the possibility of using a global model for multiple component determination of the intermediate products.

Global model validation

Specificity

Specificity refers to the ability of the method to determine the analyte without interference in the presence of other possible components. It usually adopts the method by comparing characteristic absorption peaks. However, it is difficult to perform a specific analysis of the spectrum, as the UV spectrum contains many poor specificities and low-resolution information about the sample for identifying the mixtures. Principal component analysis (PCA) can be used to extract the main features of UV spectral data and evaluate whether the mixture could be analyzed simultaneously without interference^[35–36]. The PCA scores PC1 to PC3 of UV spectra for standard substances are shown in Figure 4. The six components could be distinguished, which shows that the method is specific for the components and that the components will not interfere with each other in the analysis process.

Trueness

The trueness of an analytical method refers to the closeness of agreement between the average of the measured results and the reference value^[24,37]. The term relative bias is commonly used as the measure of trueness. The calculated values over the studied concentration levels are presented in Table 4. The relative bias of the total sugar content model in all concentration ranges was more than 15%, and the relative bias of other components was within $\pm 15\%$, except for DSS, which had a concentration of 0.0028 mg/g, and HSYA, which had a concentration of 0.0030 mg/g. The results showed that there was a large systematic deviation, and the relative bias values were larger if the target was at a relatively lower concentration. This might be because the sample concentration is close to the UV technology detection limit.

Precision

The precision is an indicator of the random error of measurement. To evaluate intra- and interprocess variance, the precision parameters of repeatability and intermediate precision were determined and are presented in Table 4. Similar to that with relative bias, the method precision decreased as the concentration decreased. However, the repeatability was close to the intermediate precision, showing favorable reproducibility of the method.

Table 3

The results for the global models of Danhong intermediates.

Index	Numbers of factors	Wavelength range (nm)	Preprocessing method	RMSECV ($\mu\text{g/g}$)	RMSEP ($\mu\text{g/g}$)	R^2_{cal}	R^2_{pre}
DSS	5	216–600	Raw	1.16	0.61	0.974	0.970
PA	3	190–600	1st derivative	0.43	0.17	0.988	0.997
RA	4	216–600	Raw	0.32	0.15	0.965	0.995
SAA	5	216–600	1st derivative	2.55	0.93	0.987	0.989
SAB	3	216–600	2nd derivative	2.93	1.75	0.998	0.995
HSYA	3	216–600	Raw	0.61	0.35	0.996	0.998
Total sugar	4	190–600	SG filter	87.27	26.07	0.970	0.991
Total phenolic	3	190–600	1st derivative	4.26	2.69	0.993	0.994

DSS: Danshensu; HSYA: Hydroxyl safflor yellow A; PA: Protocatechualdehyde; RA: Rosmarinic acid; RMSECV: Root-mean-square error of cross-validation; RMSEP: Root-mean-square error of the prediction set; SAA: salvianolic acid A; SAB: salvianolic acid B.

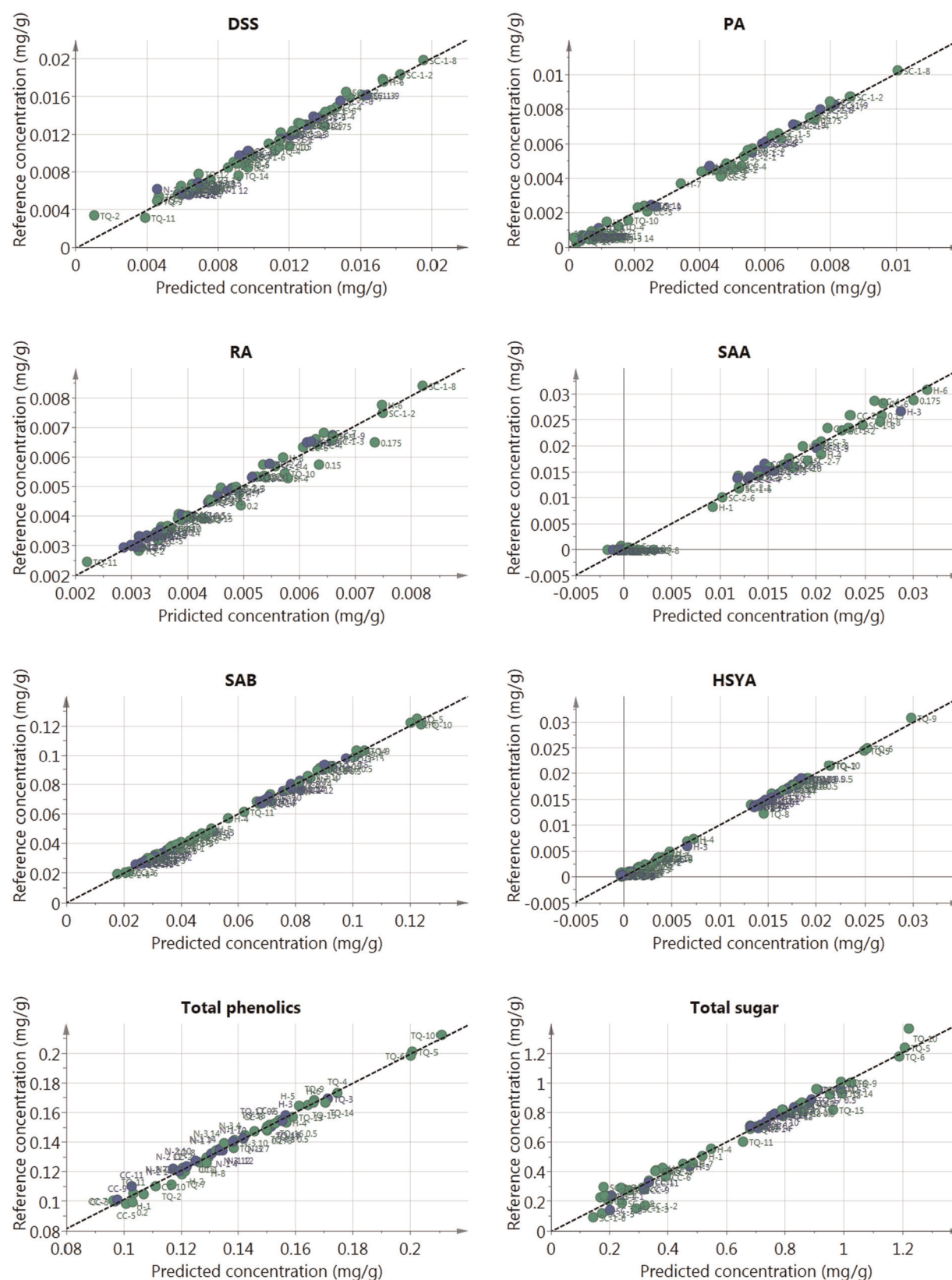


Figure 3. The correlation plots between the predicted values and the reference values (green circles stand for calibration sets, while blue circles stand for validation sets).

Accuracy

Accuracy is used to evaluate the total error of the method, including systematic error and random error. A probability of 95% was employed for the β -ETI, whereas acceptance limits λ of $\pm 15\%$ were considered. Table 4

provides the computed relative 95% β -expectation tolerance limits for the studied concentration levels. The accuracy profiles of the developed models are illustrated in Figure 5. This showed that all samples of PA (Figure 5B), RA (Figure 5C), and SAB (Figure 5E) had

Table 4
Validation results of the developed models for the intermediate samples.

124

Table 4
(continued).

Analyte	Concentration (mg/g)	Trueness		Precision		Accuracy Relative β -expectation tolerance limits (%)
		Recovery (%)	Relative bias (%)	Repeatability (RSD%)	Intermediate precision (RSD%)	
HSYA	0.0153	95.42	-4.58	0.68	1.77	[-9.66, 0.50]
	0.0116	94.15	-5.85	0.27	3.25	[-16.40, 4.70]
	0.0091	90.83	-9.17	0.70	2.10	[-15.38, -2.97]
	0.0070	89.13	-10.87	0.24	2.42	[-18.68, -3.06]
	0.0059	85.19	-14.81	1.00	3.68	[-26.03, -3.58]
	0.0030	67.54	-32.46	1.44	2.88	[-40.07, -24.84]
Total phenolic	0.2131	95.87	-4.13	0.10	0.80	[-6.71, -1.55]
	0.1733	93.76	-6.24	0.43	0.63	[-7.70, -4.78]
	0.1371	92.71	-7.29	0.50	1.13	[-10.40, -4.19]
	0.1038	96.01	-3.99	0.27	1.02	[-7.11, -0.88]
	0.0885	94.44	-5.56	0.68	1.80	[-10.74, -0.38]
	0.0448	92.81	-7.19	0.85	2.12	[-13.19, -1.19]
Total sugar	1.3203	74.14	-25.86	0.30	1.40	[-30.23, -21.49]
	1.0611	73.65	-26.35	0.55	1.05	[-29.10, -23.60]
	0.8007	76.20	-23.80	0.61	1.85	[-29.27, -18.34]
	0.6157	78.32	-21.68	0.36	2.11	[-28.37, -14.98]
	0.5165	78.92	-21.08	0.73	1.98	[-26.80, -15.35]
	0.2599	83.17	-16.83	1.90	2.70	[-23.00, -10.67]

DSS: Danshensu; HSYA: Hydroxyl safflor yellow A; PA: Protocatechualdehyde; RA: Rosmarinic acid; SAA: salvianolic acid A; SAB: salvianolic acid B.

an acceptable β -ETI. The β -ETI of DSS exceeded the acceptance limits in a wide range, but since the overall concentration was very low, the developed method was considered valid when the concentration of DSS is greater than 0.014 mg/g (Figure 5A). For SAA, the systematic error of the full concentration range was considerably close to 0 (Figure 5D), whereas the fluctuations in predicted values presented poor repeatability. Owing to the easy degradation of SAA, it can be speculated that the time interval was inconsistent between the HPLC determination and the spectra scan, and thus, the model construction might produce a certain deviation. The accuracy of the HSYA model did not fulfill predefined acceptance criteria at the lower concentration levels (Figure 5F). There was a large systematic error in the model of total sugar but satisfactory repeatability, presumably because most of the sugars have no UV absorption, and thus, it is not suitable to use UV spectroscopy for modeling. Except for the model of total sugar, the predicted values and references of DSS, PA, RA, SAA, SAB, HSYA, and total phenolic content all passed the Mann-Whitney test, which explains why there was no significant difference in the values measured by the two methods.

Linearity

The linearity of an analytical method is the ability to obtain results directly proportional to the amount of the analyte in the sample within a given range^[15]. Figure 6 depicts the linear profile of the eight developed models together with R^2 and intercept and slope values of the linear fit. A higher R^2 , which was greater than 0.9, indicated that there was a high correlation between the predicted value and the measured value. The linearity of the PA, RA, SAB, and total phenol models was good in the

entire concentration range; for the HSYA model, linearity was satisfied when the concentration was higher than 0.01 mg/g. The reasons for the poor linearity of the DSS, SAA, and total sugar content models were similar to those of the accuracy.

Robustness

The robustness of an analytical method refers to its capacity to remain unaffected by deliberate variations in the method parameters during normal usage. Since the developed UV methods are off-line measurements, variation in samples was introduced to test the robustness of the models rather than changing the temperature and other process conditions. The term RPD is used for evaluation. Standard deviation, RMSEP, and RPD values are shown in Table 5. The RPD value ($RPD > 8$) indicated the excellent performance of the models (PA and RA) and that the models could be used for any application. If $RPD > 5$, the robustness of the models is acceptable (SAB, HSYA, and total phenolic). For total sugar, $RPD < 2$ indicated that the robustness of the model was not good.

Conclusion

In this study, simple, rapid, and nondestructive global models based on UV spectroscopy with adaptive sample preparation were constructed to determine eight components in Danhong intermediate products, including the extract, concentration, and ethanol precipitation solution, among others. The high R^2 and low RMSEP showed the satisfactory performance of the models. The method was externally validated based on specificity, trueness, precision, accuracy, linearity, and robustness. The results showed that five components (PA, RA, SAB, HSYA, and

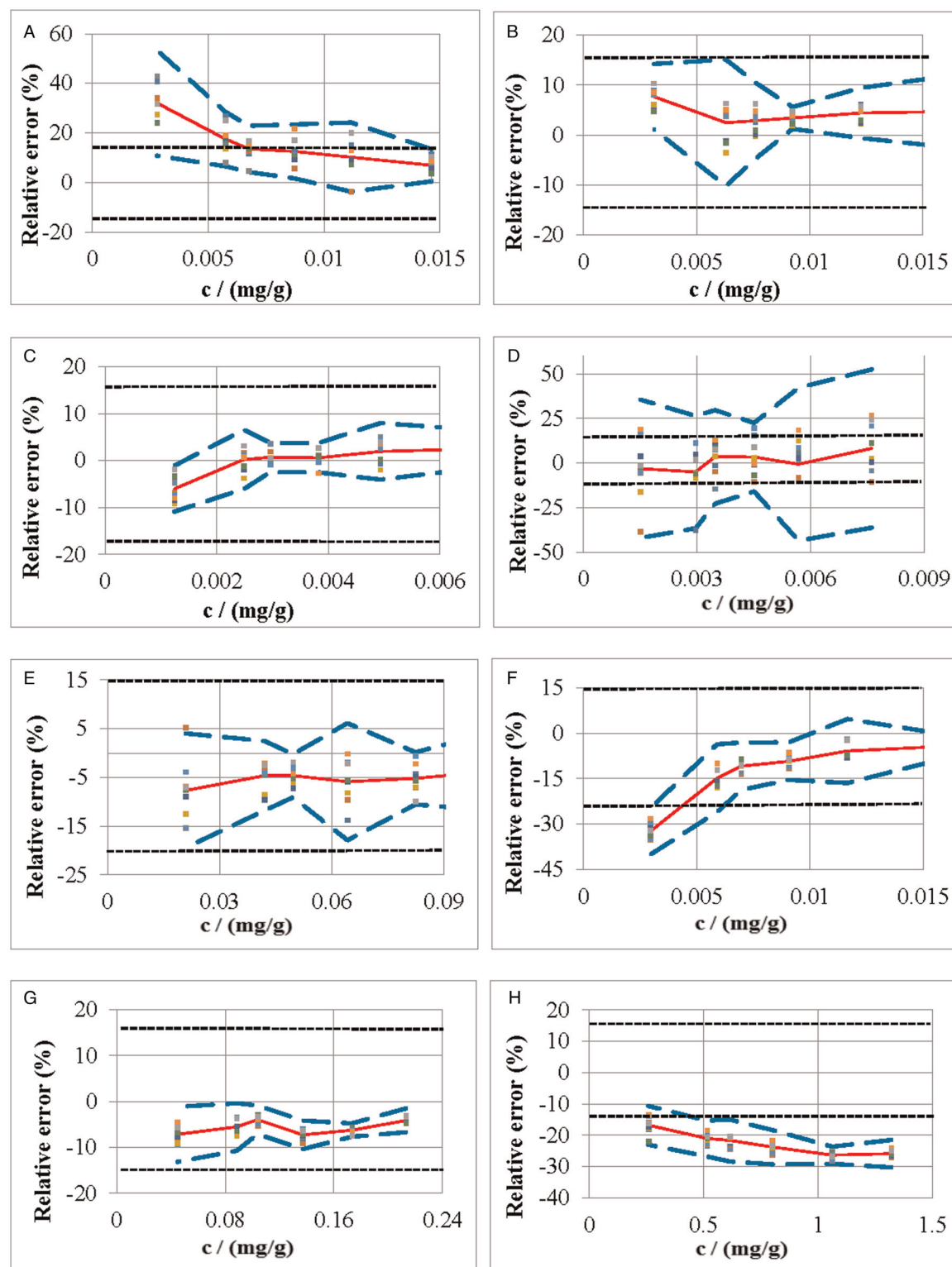


Figure 5. The accuracy profiles for the global models. (A) DSS; (B) PA; (C) RA; (D) SAA; (E) SAB; (F) HSYA; (G) total phenolic content; (H) total sugar content. The plain red line is the relative bias, the dashed blue lines are the β -expectation tolerance limits ($\beta=95\%$), and the dotted black lines are the acceptance limits ($\pm 15\%$). DSS: Danshensu; HSYA: Hydroxyl safflor yellow A; PA: Protocatechualdehyde; RA: Rosmarinic acid; SAA: salvanolic acid A; SAB: salvanolic acid B.

total phenolic content) of the models passed all validation tests and were suitable for quantitative determination. The DSS and SAA models were found to be appropriate depending on the concentration of samples and the acceptance limit. The unsatisfactory performance of the total sugar model indicates that the component is not

suitable for modeling with UV or that does not apply to the entire process. This study is the first to use the accuracy profile to comprehensively validate global models. The modeling method with high efficiency provides a reference for detection during complex Chinese medicine pharmaceutical processes.

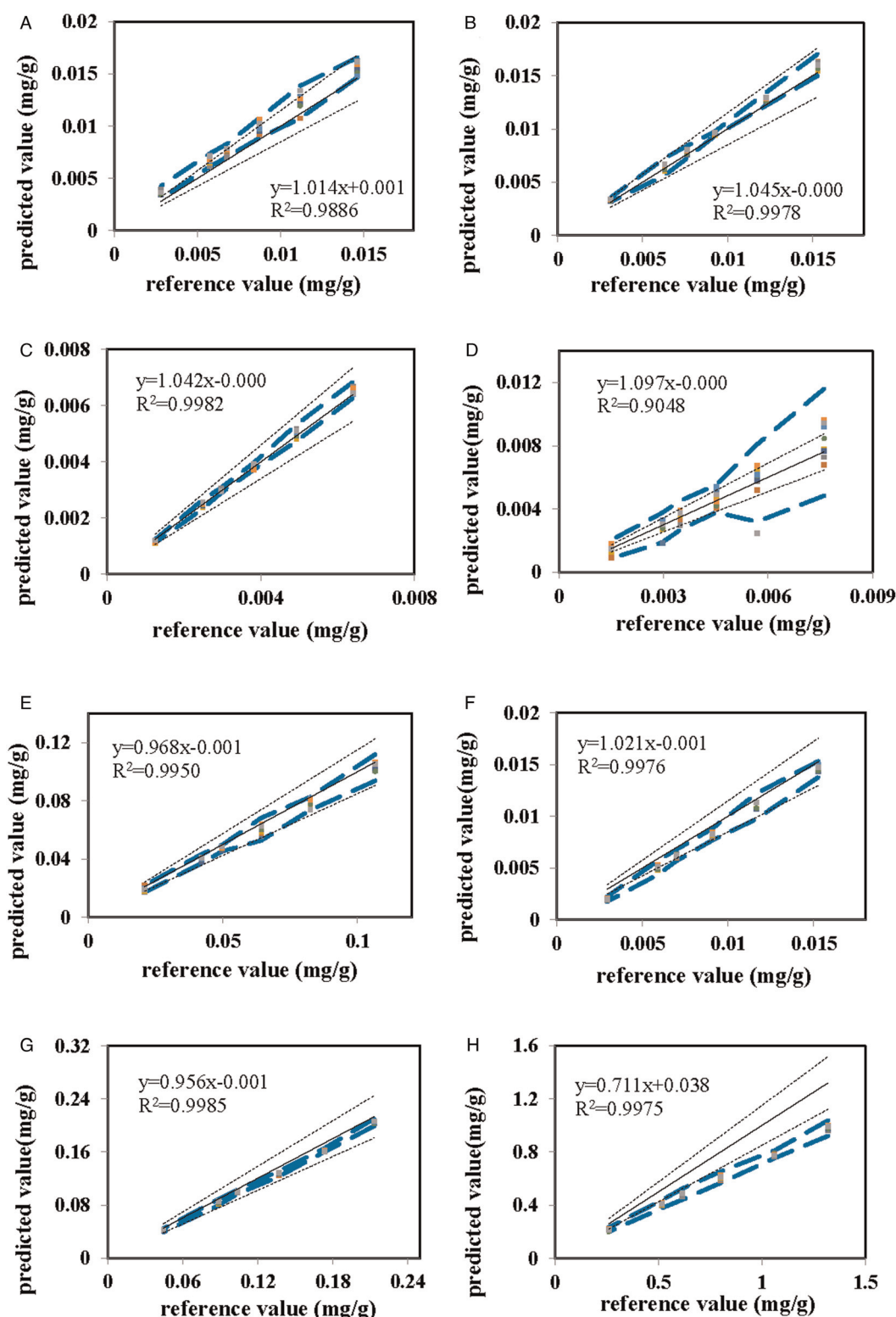


Figure 6. The linearity of each index component. (A) DSS; (B) PA; (C) RA; (D) SAA; (E) SAB; (F) HSYA; (G) total phenolic content; (H) total sugar content. DSS: Danshensu; HSYA: Hydroxyl safflor yellow A; PA: Protocatechualdehyde; RA: Rosmarinic acid; SAA: salvianolic acid A; SAB: salvianolic acid B.

Conflict of interest statement

None.

Funding

This work was supported by the National S & T Major Project of China (2018ZX09201011-002).

Author contributions

Jie Zhao: Participated in writing (original draft) the paper and data analysis; Zimei Zhou: Participated in writing (original draft) the paper and performing research and data analysis; Fang Zhao: Participated in writing (review & editing) the paper; Xu Yan: Participated in research design; Jianyang Pan: Contributed new reagents or

Table 5
The SD, RMSEP, and RPD values for each index component.

Analyte	SD (mg/g)	RMSEP (mg/g)	RPD
DSS	0.0039	0.0011	3.60
PA	0.0042	0.0005	9.29
RA	0.0018	0.0001	16.08
SAB	0.0273	0.0036	7.55
SAA	0.0023	0.0007	3.42
HSYA	0.0041	0.0008	4.96
Total phenolic	0.0535	0.0077	6.96
Total sugar	0.2517	0.2097	1.20

DSS: Danshensu; HSYA: Hydroxyl safflor yellow A; PA: Protocatechualdehyde; RA: Rosmarinic acid; RMSEP: Root-mean-square error of the prediction set; RPD: Residual prediction deviation; SAA: salvianolic acid A; SAB: salvianolic acid B.

analytic tools; Haibin Qu: Participated in conceptualization, research design, and writing (review & editing) the paper.

Ethical approval of studies and informed consent

Not applicable.

Acknowledgements

None.

References

- [1] Chen GH, Sun FR, Wang SG, et al. Enhanced extraction of essential oil from *Cinnamomum cassia* bark by ultrasound assisted hydrodistillation. *Chin J Chem Eng* 2021;36(8):38–46.
- [2] Li WZ, Zhao F, Pan JY, et al. Influence of ethanol concentration of extraction solvent on metabolite profiling for *Salviae Miltiorrhizae Radix et Rhizoma* extract by ¹H NMR spectroscopy and multivariate data analysis. *Process Biochem* 2020;97:158–167.
- [3] Tian G, Wei YD, Zhao J, et al. Application of pulsed spray and moisture content control strategies on quality consistency control in fluidized bed granulation: a comparative study. *Powder Technol* 2020;363(C):232–244.
- [4] FDA Guidance for Industry: PAT-A Framework for Innovative Pharmaceutical Development. 2004.
- [5] Pinto ASS, Ribeiro MPA, Farinas CS. Fast spectroscopic monitoring of inhibitors in the 2G ethanol process. *Bioresour Technol* 2018;250:148–154.
- [6] Zhao J, Tian G, Qiu YY, et al. Rapid quantification of active pharmaceutical ingredient for sugar-free Yangwei granules in commercial production using FT-NIR spectroscopy based on machine learning techniques. *Spectrochim Acta A Mol Biomol Spectrosc* 2021;245:118878.
- [7] Huang HX, Lyu YY, Sun XY, et al. Rapid determination of tannin in Danshen and Guanxinning injections using UV spectrophotometry for quality control. *J Innov Opt Health Sci* 2018;11:1850034.
- [8] Liu XS, Wu ZZ, Yang K, et al. Quantitative analysis combined with chromatographic fingerprint for comprehensive evaluation of Danhong injection using HPLC-DAD. *J Pharm Biomed Anal* 2013;76:70–74.
- [9] Li WL, Qu HB. Rapid quantification of phenolic acids in *Radix Salviae Miltiorrhiza* extract solutions by FT-NIR spectroscopy in transreflective mode. *J Pharm Biomed Anal* 2010;52:425–431.
- [10] Song XC, Canellas E, Asensio E, et al. Predicting the antioxidant capacity and total phenolic content of bearberry leaves by data fusion of UV-vis spectroscopy and UHPLC/Q-TOF-MS. *Talanta* 2020;213:120831.
- [11] Zhao J, Li WL, Qu HB, et al. Real-time monitoring and fault detection of pulsed-spray fluid-bed granulation using near-infrared spectroscopy and multivariate process trajectories. *Particuology* 2020;53:112–123.

- [12] Cardoso VGK, Poppi RJ. Non-invasive identification of commercial green tea blends using NIR spectroscopy and support vector machine. *Microchem J* 2021;164:106052.
- [13] Clavaud M, Roggo Y, Dégardin K, et al. Global regression model for moisture content determination using near-infrared spectroscopy. *Eur J Pharm Biopharm* 2017;119:343–352.
- [14] Hodge GR, Acosta JJ, Unda F, et al. Global near infrared spectroscopy models to predict wood chemical properties of *Eucalyptus*. *J Near Infrared Spec* 2018;26:117–132.
- [15] International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. vol Q2. (R1): Validation of Analytical Procedures: Text and Methodology. 2005.
- [16] Marzouk HM, Teaima MH, El-Desoky AM, et al. A validated HPLC-DAD method for simultaneous determination of major bioactive constituents of antidiabetic herbal extracts; application to a newly co-formulated nutraceutical product. *Microchem J* 2020;159:105472.
- [17] Brouckaert D, Uyttersprot JS, Broeckx W, et al. Calibration transfer of a Raman spectroscopic quantification method for the assessment of liquid detergent compositions from at-line laboratory to in-line industrial scale. *Talanta* 2018;179:386–392.
- [18] Hubert P, Nguyen-Huu JJ, Boulanger B, et al. Harmonization of strategies for the validation of quantitative analytical procedures—a SFSTP proposal – Part I. *J Pharm Biomed Anal* 2004;36:579–586.
- [19] Hubert P, Nguyen-Huu JJ, Boulanger B, et al. Harmonization of strategies for the validation of quantitative analytical procedures—a SFSTP proposal – Part II. *J Pharm Biomed Anal* 2007;45:70–81.
- [20] Hubert P, Nguyen-Huu JJ, Boulanger B, et al. Harmonization of strategies for the validation of quantitative analytical procedures—a SFSTP proposal – Part III. *J Pharm Biomed Anal* 2007;45:82–96.
- [21] Hubert P, Nguyen-Huu JJ, Boulanger B, et al. Harmonization of strategies for the validation of quantitative analytical procedures: a SFSTP proposal – Part IV. Examples of application. *J Pharm Biomed Anal* 2008;48:760–771.
- [22] Tighrine A, Amir Y, Mamou M. Application of an accuracy profile strategy based on the beta-expectation tolerance interval for the validation of a liquid chromatography analytical method for the quantification of benzoic acid and its salts in different foodstuffs. *J AOAC Int* 2017;100:241–244.
- [23] Mantanus J, Ziémons E, Lebrun P, et al. Active content determination of non-coated pharmaceutical pellets by near infrared spectroscopy: method development, validation and reliability evaluation. *Talanta* 2010;80:1750–1757.
- [24] Kauppinen A, Toivainen M, Lehtonen M, et al. Validation of a multipoint near-infrared spectroscopy method for in-line moisture content analysis during freeze-drying. *J Pharm Biomed Anal* 2014;95:229–237.
- [25] Pereira LSA, Carneiro MF, Botelho BG, et al. Calibration transfer from powder mixtures to intact tablets: a new use in pharmaceutical analysis for a known tool. *Talanta* 2016;147:351–357.
- [26] Bostijn N, Hellings M, Van der Veen M, et al. In-line UV spectroscopy for the quantification of low-dose active ingredients during the manufacturing of pharmaceutical semi-solid and liquid formulations. *Anal Chim Acta* 2018;1013:54–62.
- [27] Bourdon F, Lecoer M, Odou P, et al. Complementarity of UV-PLS and HPLC for the simultaneous evaluation of antiemetic drugs. *Talanta* 2014;120:274–282.
- [28] Yi XJ, Zhu JF, Zhang JH, et al. Investigation of the reverse effect of Danhong injection on doxorubicin-induced cardiotoxicity in H9c2 cells: insight by LC-MS based non-targeted metabolomic analysis. *J Pharm Biomed Anal* 2018;152:264–270.
- [29] Zhu JF, Yi XJ, Huang P, et al. Drug-protein binding of Danhong injection and the potential influence of drug combination with aspirin: insight by ultrafiltration LC-MS and molecular modeling. *J Pharm Biomed Anal* 2017;134:100–107.
- [30] Tian G, Wei YD, Zhao J, et al. Application of near-infrared spectroscopy combined with design of experiments for process development of the pulsed spray fluid bed granulation process. *Powder Technol* 2018;339:521–533.
- [31] Commission CP. Pharmacopoeia of the People's Republic of China. vol. IV. 2015;China Medical Science Press.
- [32] Cuesta G, Suarez N, Bessio MI, et al. Quantitative determination of pneumococcal capsular polysaccharide serotype 14 using a modification of phenol-sulfuric acid method. *J Microbiol Methods* 2003;52:69–73.
- [33] Kennard RW, Stone LA. Computer aided design of experiments. *Technometrics* 1969;11(1):137–148.

- [34] Zacharis CK, Vastardi E. Application of analytical quality by design principles for the determination of alkyl p-toluenesulfonates impurities in Aprepitant by HPLC. Validation using total-error concept. *J Pharm Biomed Anal* 2018;150:152–161.
- [35] Luoma P, Golabgir A, Brandstetter M, et al. Workflow for multi-analyte bioprocess monitoring demonstrated on inline NIR spectroscopy of *P. chrysogenum* fermentation. *Anal Bioanal Chem* 2017;409:797–805.
- [36] Luo XF, Yu X, Wu XM, et al. Rapid determination of Paeoniae Radix using near infrared spectroscopy. *Microchem J* 2008;90:8–12.
- [37] Kamal AH, Marie AA, Hammad SF. Validated spectrophotometric methods for simultaneous determination of nebiivolol hydrochloride and valsartan in their tablet. *Microchem J* 2020;155:104741.

How to cite this article: Zhao J, Zhou ZM, Zhao F, Yan X, Pan JY, Qu HB. Development and validation of global prediction models for monitoring the manufacturing process of herbal medicine by ultraviolet spectroscopy. *Acupunct Herb Med* 2022;2(2):118–129. doi: 10.1097/HM9.0000000000000034