Original Article

Acupuncture and Herbal Medicine

OPEN

In-line spectroscopy combined with multivariate analysis methods for endpoint determination in column chromatographic adsorption processes for herbal medicine

Cheng Jiang[1,2], Haibin Qu[1,3,*]

Abstract

Objective: In a chromatographic cycle, the adsorption process is a critical unit operation that has a significant impact on downstream processes and, ultimately, the quality of the final products. The development of a rapid method to determine the endpoints of adsorption processes in a large-scale manufacturing is of substantial importance for herbal medicine (HM) manufacturers.

Methods: In this study, the adsorption of saponins on a macroporous resin column chromatograph, a critical unit operation in *Panax notoginseng* (Burkill) F.H.Chen injection manufacturing, was considered as an example. The evaluation results of in-line ultraviolet and visible spectra combined with various multivariate analysis methods, including the moving block standard deviation (MBSD), difference between the moving block average and the target spectrum (DMBA-TS), soft-independent modeling of class analogy (SIMCA), and partial least-squares discriminant analysis (PLS-DA), were compared.

Results: MBSD was unsuitable for adsorption processes. The relative standard errors of prediction between the predicted and experimental endpoints were 13.2%, 4.67%, and 5.71% using DMBA-TS, SIMCA, and PLS-DA, respectively.

Conclusions: Among the considered analysis methods, SIMCA and PLS-DA were more effective for endpoint determination. The results of this study provide a more comprehensive overview of the effectiveness of various multivariate analysis methods to facilitate the selection of the most suitable method. This study was also conducive to address the issues of the in-line detection of adsorption endpoints to guide practical HM manufacturing.

Keywords: Adsorption, Endpoint, Herbal medicine, Multivariate analysis, Ultraviolet and visible

Graphical abstract: http://links.lww.com/AHM/A23

Introduction

Column chromatography has a wide range of applications in many fields, including herbal medicine (HM) manufacturing, biological medicine manufacturing, and wastewater treatment. During the manufacturing of HM, especially botanical injections, column chromatography plays an important role in improving the safety and efficacy of the final products. It has been used to separate several types of bioactive substances from medicinal herbs, including saponins, polyphenols, flavonoids, and alkaloids[1]. Each chromatographic cycle typically consists of adsorption, washing, and elution. The adsorption process is a critical unit operation and may have a significant impact on the downstream processes and quality of the final products. Insufficient adsorption time may result in a low efficiency of adsorbent utilization, whereas excessive adsorption time may result in high variability in the final product quality and low efficiency of solute adsorption[2]. The determination of the optimal adsorption endpoint is of considerable importance in a large-scale manufacturing.

Breakthrough time, which is defined as the time required for a certain number of active components to be detected in the effluent, is generally considered a suitable endpoint[3]. However, several factors may influence the breakthrough time, such as variations in loading solutions, fluctuations in operating parameters, and variations in the adsorption capacity of the adsorbents. Consequently, traditional HM manufacturing, which is conducted with a fixed adsorption time as the endpoint, cannot adapt to changes in the above-mentioned factors. The breakthrough time can be determined by an offline assessment of the active components in the effluents *via*

[1] *College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, China;* [2] *Zhejiang Provincial Key Laboratory of Cancer Prevention and Treatment Technology of Integrated Traditional Chinese and Western Medicine, Zhejiang Academy of Traditional Chinese Medicine, Tongde Hospital of Zhejiang Province, Hangzhou, China;* [3] *Innovation 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.*

Acupuncture and Herbal Medicine (2022) 2:4

Received 15 January 2022 / Accepted 11 July 2022

http://dx.doi.org/10.1097/HM9.0000000000000035

chromatographic techniques. However, the long duration required for analysis may ultimately delay critical manufacturing decisions. Therefore, the development of a rapid method to determine the endpoints of adsorption processes is of critical importance for HM manufacturers.

Spectroscopy is an effective choice for process analysis because it requires no sample preparation and can provide real-time information about the processes. Thus far, several studies have used spectroscopic techniques for endpoint determination in pharmaceutical processes, including blending, fluidized drying, fluidized bed granulation, coating, extraction, and chemical synthesis[4-9]. Spectroscopic techniques have also been reported to monitor the column chromatographic elution processes of *Coptis chinensis* Franch[10] and *Salvia miltiorrhiza* Bunge[11]. Ultraviolet and visible (UV-VIS) spectroscopy is a form of electron spectroscopy that is based on the excitation of electrons[12]. It possesses the advantages of high sensitivity and inexpensive instrumentation. However, the performance of in-line UV-VIS spectroscopy in the determination of adsorption endpoints during HM manufacturing is not yet explored.

Numerous qualitative methods have been applied for process monitoring and can be classified into several categories. The first consists of calibration-free methods, including moving block standard deviation (MBSD) and monitoring specific spectral peak intensity of active substances[13-14]; these approaches are relatively straightforward and require no additional information for modeling. The second category relies on methods that estimate the difference between the spectra obtained during the process and the target spectrum, including the difference between the moving block average and the target spectrum (DMBA-TS) and spectra linear superposition method[13,15]. The third category consists of chemometric methods based on principal component analysis (PCA), including the soft-independent modeling of class analogy (SIMCA) and principal component score distance analysis (PC-SDA)[16-18]. The fourth category consists of chemometric methods based on a classical partial least squares regression (PLSR) where the response variable is a categorical variable, including partial least-squares discriminant analysis (PLS-DA)[19-20]. However, no critical study has been conducted to benchmark the potential of various multivariate analysis methods for HM adsorption processes thus far. Therefore, we conducted a comparative study to evaluate the effectiveness of multivariate analysis methods and facilitate the selection of the most appropriate method for the in-line detection of adsorption endpoints during HM manufacturing.

This study compared existing multivariate analysis methods to determine adsorption endpoints using in-line UV-VIS spectra. Four representative qualitative methods, including MBSD, DMBA-TS, SIMCA, and PLS-DA, were investigated. The adsorption of saponins on a D-101 macroporous resin column chromatograph, a critical unit operation in *Panax notoginseng* (Burkill) F.H.Chen injection manufacturing, was considered as an example. This study aims to provide a comprehensive and holistic overview of the advantages and disadvantages of various multivariate analysis methods to effectively facilitate their application in guiding practical HM adsorption processes.

Materials and Methods

Samples, chemicals, and reagents

The medicinal herb *Panax notoginseng* (Burkill) F.H.Chen was purchased from the Wenshan Panax Notoginseng International Trade Center (Yunnan, China). The standard compounds notoginsenoside R_1, ginsenoside Rg_1, ginsenoside Re, ginsenoside Rb_1, and ginsenoside Rd were purchased from Ronghe Pharmaceutical Technology Co., Ltd. (Shanghai, China) with purities of over 99%. D-101 macroporous resin was purchased from Haiguang Chemical Co., Ltd. (Tianjin, China). HPLC-grade acetonitrile was purchased from Merck (Darmstadt, Germany). Ultrapurified water was prepared using a Milli-Q academic water purification system (Millipore, Billerica, MA, USA).

Process description

In this study, seven experiments were conducted. Experiments 1 to 6 were designed to establish the models, while experiment 7 evaluated the predictive capability of the models. The loading solution was prepared using several steps. Briefly, dry *Panax notoginseng* (Burkill) F.H.Chen powders with different harvest seasons (spring and winter) and medicinal components (rhizome and taproot) were refluxed twice using twice the amount of 90% ethanol-water solution (v/w); the extracts were then filtered. Thereafter, the filtrates were merged and concentrated. Subsequently, the concentrated solutions were diluted with water to different mass concentrations (0.2, 0.4, and 1.0 g/mL) and some precipitates formed. The loading solutions were obtained *via* vacuum suction filtration. The harvest season, medicinal components, and weight concentration corresponding to each loading solution are presented in Supplementary Table S1 (http://links.lww.com/AHM/A23). The details of the loading solution preparation can be found in our previous report[21].

The adsorption process in each experiment was conducted based on several procedures. D-101 macroporous resin column chromatography with a 1.5-cm inner diameter and 11-cm bed height was employed. The column was first equilibrated with five bed volumes (BVs) of water and subsequently loaded with the loading solution at various flow rates (0.47, 1.0, and 2.8 mL/min). The flow rate was controlled using a BT300-2J peristaltic pump (Longer, Hebei, China). Following adsorption, the column was regenerated with five BVs of 95% ethanol-water (v/v), water, 1 mol/L HCl solution, water, and 1 mol/L NaOH solution. A schematic of the experimental setup is depicted in Supplementary Figure S1 (http://links.lww.com/AHM/A23). The flow rates for adsorption in each experiment are listed in Supplementary Table S1 (http://links.lww.com/AHM/A23).

Throughout the adsorption process, the effluents were monitored using a T6 spectrophotometer (Purkinje, Beijing, China) integrated with the UVwin5 software. A flow cell with a 5-mm pathway was used for the in-line collection of spectra. UV-VIS spectra were collected in the region of 200 to 720 nm with a data interval of 1 nm. Ultrapure water was used as the background for the calculation of absorbance before the daily operation.

The scan speed was rapid and each spectrum could be acquired within approximately 1.7 min. The effluents were collected every 4 min using a BS-100N automatic compositional collector (Huxi, Shanghai, China) for the offline ultraperformance liquid chromatography (UPLC) analysis of notoginsenoside R_1, ginsenoside Rg_1, ginsenoside Re, ginsenoside Rb_1, and ginsenoside Rd. Details of the UPLC analytical method have been presented in a previous article[22]. All time intervals were calculated by subtracting the dead time, which is defined as the first point in time at which the potassium chloride concentration in affluent/concentration in the loading solution (C/C_0) was higher than 10% as detected by a conductivity meter[22]. Finally, the first point in time at which notoginsenoside R_1 C/C_0 was observed to be higher than 10%, as detected using UPLC, was determined to be the endpoint of each experiment[23].

Multivariate modeling

Based on the UPLC results, the in-line UV-VIS spectra were divided into two categories. One category consisted of spectra collected before the endpoint (SBEP), while the other category consisted of spectra collected after the endpoint (SAEP). PCA provides correlations between features by constructing orthogonal principal components (PCs)[24]. PCs are linear combinations of feature vectors oriented in the direction of maximum variance[24]. In this study, PCA was first performed to visualize the spectra of the two categories and to display the trajectory of the processes. The in-line UV-VIS spectra based on all wavelengths in the calibration set (Experiments 1–6) were used to establish the PCA model; the spectra in the validation set (Experiment 7) were then projected onto the model. The detailed computing formula of PCA can be found in the relevant literature[25].

Various multivariate data analysis methods, including MBSD, DMBA-TS, SIMCA, and PLS-DA, were then investigated for endpoint determination. MBSD calculates the SD for consecutive blocks of spectra over time. A time point with an SD lower than a certain threshold is usually defined as the predicted endpoint. In this study, three consecutive spectra based on all wavelengths were used, while one spectrum was moved at a time. The detailed computing formula for MBSD can be found in the relevant reference[13].

DMBA-TS calculates the difference between the spectra recorded during the process and the target spectrum. In this study, the target spectrum was defined as the average of the first spectra collected after the endpoint (FSAEP) in the calibration set. A block of three consecutive spectra recorded during the adsorption process was utilized. The time point with the minimum SD was defined as the predicted endpoint for each experiment, wherein one spectrum was moved at a time. The detailed computing formula for DMBA-TS can be found in the relevant reference[13].

SIMCA is a well-known pattern-recognition method based on PCA. It describes each class separately in PC space and builds a distinct confidence region around each class[26]. In this study, the SIMCA model was established based on two spectral classes (SBEP and SAEP) in the calibration set. The first time point with the distance from the process spectrum to the SBEP model domain higher than the critical distance (which was defined as the 95% confidence interval of distance) was defined as the predicted endpoint. The detailed computational formula for SIMCA can be found in the relevant reference[26].

PLS-DA was established based on the classical PLSR. The objective of PLS-DA is to find a model that separates classes of observations based on X variables[27]. In this study, the X matrix contained the in-line UV-VIS spectra of the effluents in the calibration set, while the Y matrix contained dummy variables that described the class membership of each spectrum; "1" was used to encode the SBEP and "0" was used to encode the SAEP. The first time point with a predicted Y less than 0.5 was defined as the predicted endpoint. The detailed computing formula for PLS-DA can be found in the relevant reference[28].

The predictive capacities of various models were evaluated by comparing the relative standard error of prediction (RSEP) between the predicted and experimental endpoints, as shown in Equation (1)[29]:

$$\mathrm{RSEP} = \sqrt{\frac{\sum \left(\hat{t}_e - t_e\right)^2}{\sum t_e^2}} \quad (1)$$

where t_e is the experimental endpoint, and $\hat{t}_e$ is the predicted endpoint. All data analyses were performed using MATLAB (version 7.5.0, MathWorks, USA) scripts written in-house.

Results

Overview of the spectral profile

The UV-VIS spectra of the effluents were collected using an inline flow cell. The UV-VIS spectra of Experiment 1 are shown in Supplementary Figure S1 (http://links.lww.com/AHM/A23). As the adsorption process progressed, the absorbance in the UV-VIS spectra gradually increased. The absorbance of several wavelengths was higher than 2, indicating that less than 1% of the incident light was transmitted. To minimize the influence of spectral noise, the absorbance at these wavelengths was set to 2. The endpoints of each experiment obtained using offline UPLC are listed in Table 1.

PCA was performed based on the in-line UV-VIS spectra in Experiments 1 to 6. The score plots (PC_1-PC_2) of the in-line UV-VIS spectra and FSAEP in each experiment are shown in Figure 1. PC_1 explained 71.6% of the variation, while PC_2 explained 18.3% of the variation. All runs followed comparable trajectories. It was demonstrated that the UV-VIS spectra could provide information on process evolution.

Endpoint determination using MBSD

MBSD was then applied to evaluate the endpoints of the adsorption processes. The SD values throughout the adsorption process for each experiment are shown in Figure 2. A gradual decrease in the SD values was observed. Endpoints of the adsorption processes were difficult to determine because the SD threshold varied with changes in the loading solutions and flow rates.

Table 1

Experimental endpoints determined using ultra-performance liquid chromatography; predicted endpoints and predicted errors using various multivariate analysis methods (min)

		Predicted endpoints using different methods			Predicted endpoint errors using different methods		
No.	Experimental endpoints	DMBA-TS	SIMCA	PLS-DA	DMBA-TS	SIMCA	PLS-DA
1	122.0	130.9	119.0	119.0	8.9	−3.0	−3.0
2	86.0	105.9	89.4	93.8	19.9	3.4	7.8
3	70.0	74.3	65.5	67.2	4.3	−4.5	−2.8
4	2.0	0.3	0.3	0.3	−1.7	−1.7	−1.7
5	17.0	21.2	14.4	21.2	4.2	−2.6	4.2
6	4.3	9.0	4.1	5.7	4.7	−0.2	1.4
7	58.0	59.2	53.9	57.4	1.2	−4.1	−0.6

DMBA-TS: difference between the moving block average and target spectrum; PLS-DA: partial least-squares discriminant analysis; SIMCA: soft-independent modeling of class analogy.

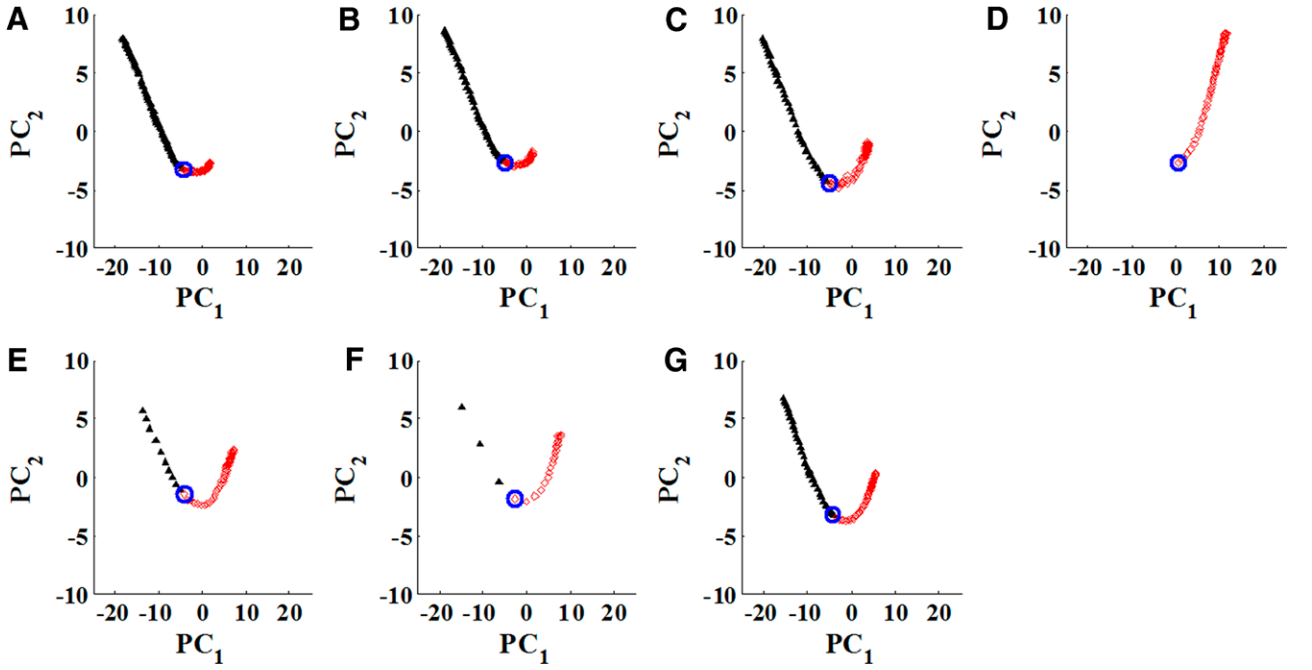

Figure 1. Scores plots of in-line ultraviolet and visible spectra by principal component analysis. (A) Experiment 1; (B) Experiment 2; (C) Experiment 3; (D) Experiment 4; (E) Experiment 5; (F) Experiment 6; (G) Experiment 7. PC_1: Principal component 1; PC_2: Principal component 2. First spectra being collected after end-point: ○; Spectra being collected before end-point: ▲; Spectra being collected after endpoint: ◇. All runs followed similar process trajectories indicated that the in-line ultraviolet and visible spectra could provide information on the process evolution.

Therefore, MBSD was not sufficiently sensitive to detect the adsorption endpoints.

Endpoint determination using DMBA-TS

Next, DMBA-TS was conducted to evaluate the endpoints. The SD values throughout the adsorption process for each experiment are shown in Figure 3. A gradual decrease in SD values was observed prior to the endpoints, whereas they gradually increased beyond the endpoints. The predicted endpoints and errors of each experiment using DMBA-TS are listed in Table 1. The RSEP between the predicted and experimental endpoints using DMBA-TS was 13.2%.

Endpoint determination using SIMCA

SIMCA was performed based on two classes of spectra. The numbers of components for the SBEP and SAEP models were four and three, respectively. The cumulative R^2 values of all extracted components were 0.996 and 0.998 for the SBEP and SAEP models, respectively. The cumulative Q^2 values for the extracted components were 0.995 and 0.998 for the SBEP and SAEP models, respectively. The distances from the spectra to the SBEP and SAEP model domains during the adsorption processes are shown in Figure 4. The SBEP and SAEP were distinguishable. The predicted endpoints and errors of each experiment using SIMCA are listed in Table 1. The RSEP between the predicted and experimental endpoints using SMICA was 4.67%.

Endpoint determination using PLS-DA

The PLS-DA model was established based on the in-line UV-VIS spectra in the calibration set. There were six components in the PLS-DA model. The cumulative R^2 value of all the extracted components was 0.998, while the cumulative Q^2 for the extracted components was 0.931. The predicted Y values for each experiment

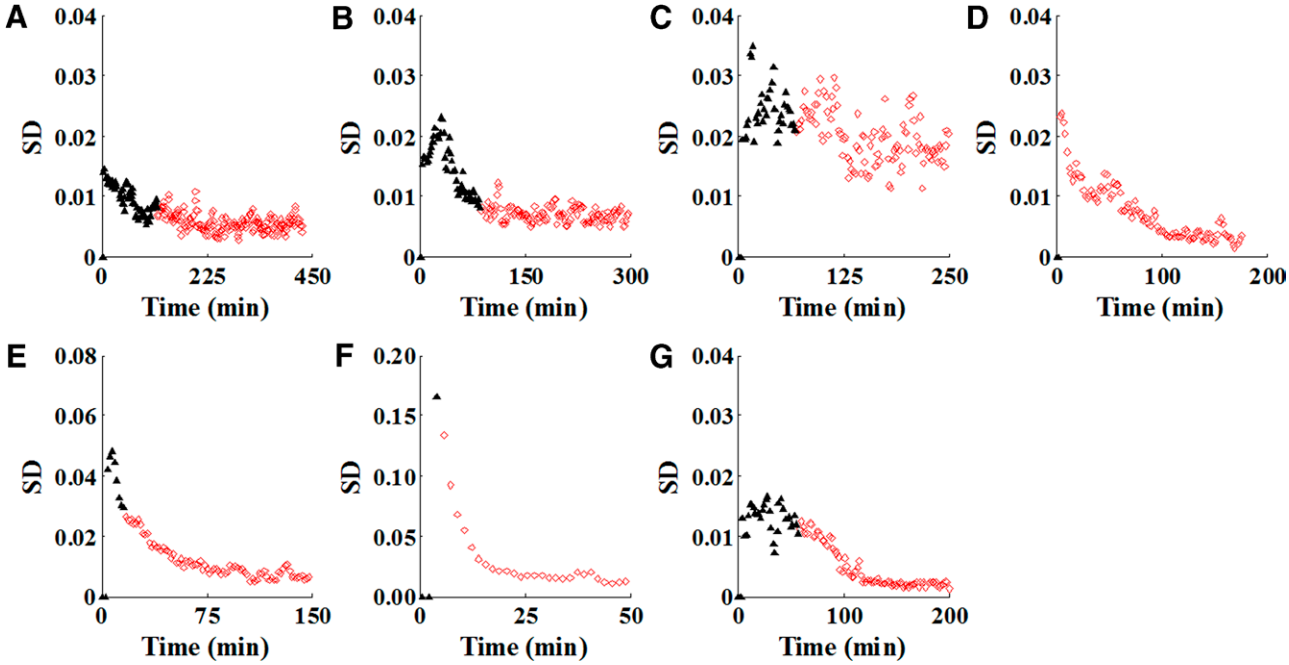

Figure 2. Endpoints determination by moving block SD. (A) Experiment 1; (B) Experiment 2; (C) Experiment 3; (D) Experiment 4; (E) Experiment 5; (F) Experiment 6; (G) Experiment 7. SD: standard deviation. Spectra being collected before endpoint: ▲; Spectra being collected after end-point: ◇. The threshold changed with the changes of loading solutions and flow rates indicated that the moving block SD was not sensitive to detect the adsorption endpoints.

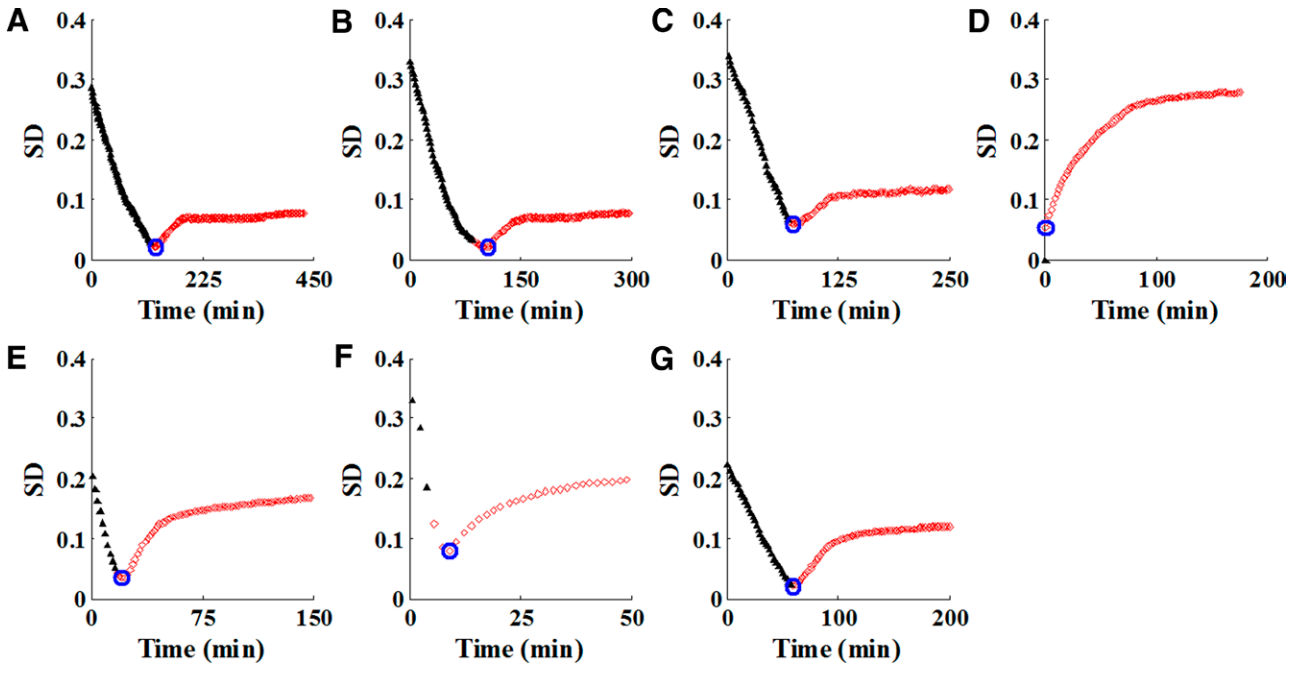

Figure 3. Endpoints determination by difference between the moving block average and the target spectrum. (A) Experiment 1; (B) Experiment 2; (C) Experiment 3; (D) Experiment 4; (E) Experiment 5; (F) Experiment 6; (G) Experiment 7. SD: standard deviation. Predicted end-point: ○; Spectra being collected before end-point: ▲; Spectra being collected after endpoint: ◇. The predicted end-points were defined to be time points with minimum SD values.

throughout the adsorption process using PLS-DA are shown in Figure 5. SBEP and SAEP could be distinguished as follows: SBEP possessed a predicted Y above 0.5, whereas SAEP possessed a predicted Y below 0.5. The predicted endpoints and errors of each experiment using PLS-DA are listed in Table 1. The RSEP between the predicted and experimental endpoints using PLS-DA was 5.71%.

Discussion

In general, adsorption reached saturation when the component concentration in the effluent was 10% of that in the loading solution[30]. Notoginsenoside R_1, ginsenoside Rg_1, ginsenoside Re, ginsenoside Rb_1, and ginsenoside Rd are the major bioactive components of the *Panax notoginseng* (Burkill) F.H.Chen injection. As revealed in

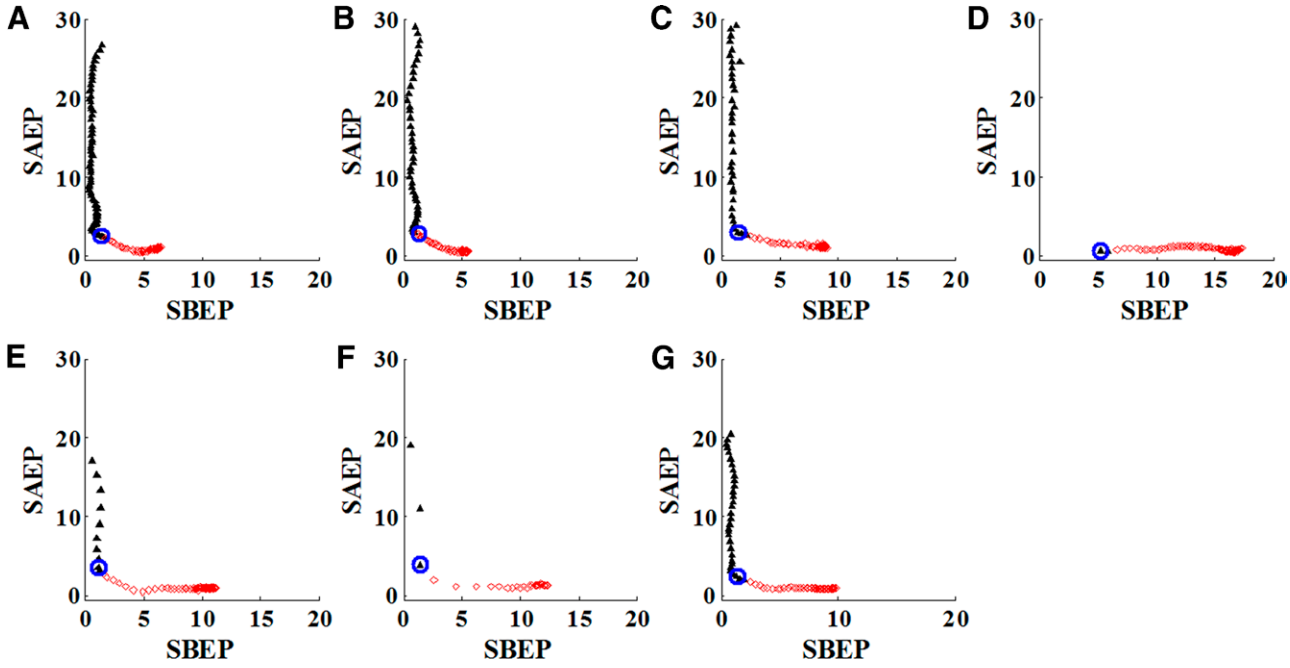

Figure 4. Endpoints determination by soft independent modeling of class analogy. (A) Experiment 1; (B) Experiment 2; (C) Experiment 3; (D) Experiment 4; (E) Experiment 5; (F) Experiment 6; (G) Experiment 7. SAEP: Spectra being collected after endpoint; SBEP: Spectra being collected before endpoint. Predicted endpoint: ○; Spectra being collected before endpoint: ▲; Spectra being collected after end-point: ◇. The predicted endpoints were defined to be the first time point with distance from process spectrum to spectra being collected before endpoint model domain outside the 95% confidence intervals of distance.

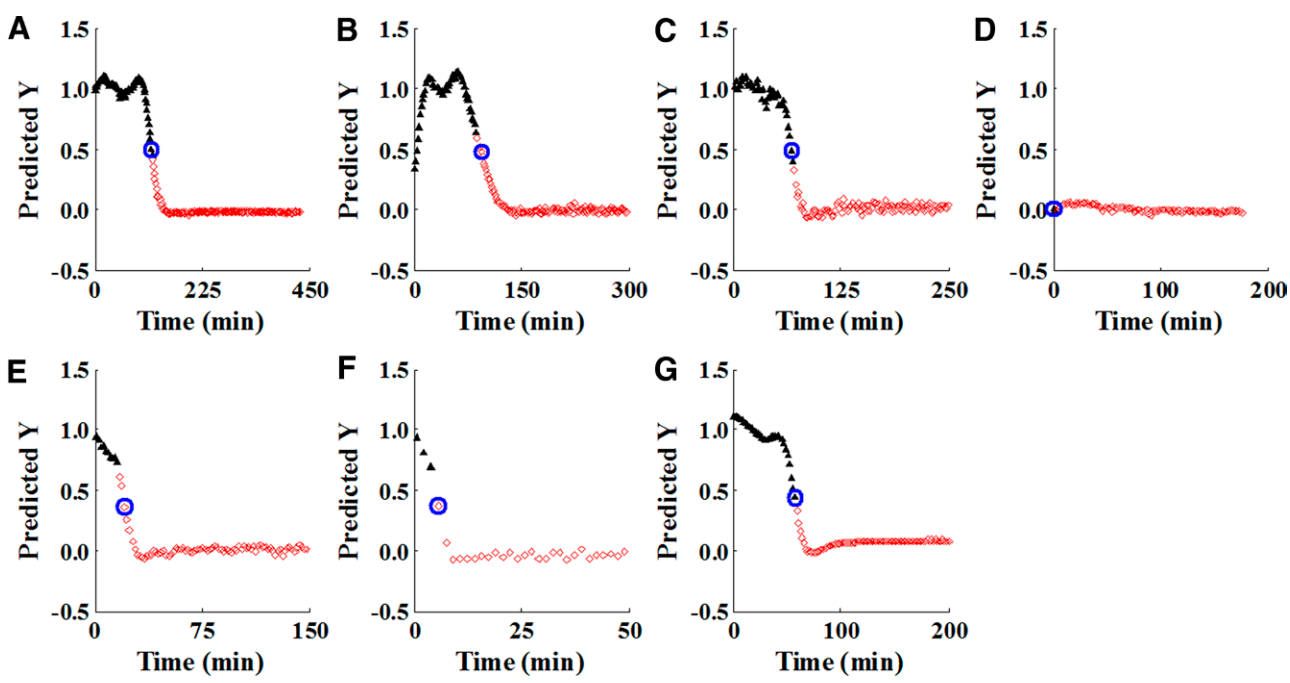

Figure 5. Endpoints determination by partial least-squares discriminant analysis. (A) Experiment 1; (B) Experiment 2; (C) Experiment 3; (D) Experiment 4; (E) Experiment 5; (F) Experiment 6; (G) Experiment 7. Predicted endpoint: ○; Spectra being collected before endpoint: ▲; Spectra being collected after endpoint: ◇. The predicted endpoints were defined to be the first time point with predicted Y below 0.5.

our previous study, notoginsenoside R_1 is easily lost in the effluent of these five saponins[23]. Therefore, it was used as the indicator component for endpoint determination. The first time-point with notoginsenoside R_1 C/C_0 higher than 10% to be detected through UPLC was determined to be the experimental endpoint.

The quality of the loading solution was dependent on the harvest season, medicinal components, and dilution ratio. When the flow rates or loading solutions were altered, both the endpoints of the adsorption processes and composition of the effluent changed. To build multivariate analysis models with wide applicability, the spectra included in the calibration set were obtained from Experiments 1 to 6, which were conducted with various flow rates, harvest seasons, medicinal components, and dilution ratios. In Experiment 4, saponins leaked rapidly from the macroporous resin. Compared with other adsorption processes, the weight

concentration of the loading solution diluent in Experiment 4 was extremely high. These results indicate that a high concentration of saponins and limited number of active sites may result in the incomplete adsorption of saponins on the macroporous resin.

MBSD was found to be unsuitable for adsorption processes. This may be explained by the fact that MBSD is particularly well-suited for processes with a gradual increase or decrease in the components. Prior to the endpoint of adsorption, no saponins were detected in the effluent. However, components such as pigments, whose adsorption capacity on the resin is weaker than that of saponins, may be detected in the effluent. Additionally, the threshold was extremely high in Experiments 5 and 6. Compared with the other adsorption processes, the flow rates in Experiments 5 and 6 were high. The results indicate that a high flow rate may result in a rapid change in the spectra as a consequence high threshold of MBSD. Although MBSD has been widely recommended for endpoint determination in the blending processes of chemical medicine, it is incompatible with the adsorption process of HM.

The RSEPs between the predicted and experimental endpoints were 13.2%, 4.67%, and 5.71% using DMBA-TS, SIMCA, and PLS-DA, respectively. Compared with MBSD, DMBA-TS can be used to assess the closeness of the process spectrum to the target spectrum; however, the description of the target spectrum is required prior to modeling. The RSEPs obtained using SIMCA and PLS-DA were smaller than those obtained *via* DMBA-TS, indicating that SIMCA and PLS-DA were more suitable for endpoint determination in adsorption processes. A possible reason might be that the MBSD was a calibration-free method, whereas the DMBA-TS model was established based on the target spectrum. As shown in Figure 1, the in-line UV-VIS spectra during the adsorption processes can also provide information on process evolution. Therefore, SIMCA and PLS-DA, which were modeled based on all spectra in the calibration set, contained more information. Moreover, models using both MBSD and DMBA-TS were established based on all wavelengths in the UV-VIS spectral region. SIMCA and PLS-DA are representative methods for information extraction, which are helpful for classifying large amounts of data. It was demonstrated that models containing more information and processing with appropriate information extraction enabled the favorable detection of endpoints. SIMCA and PLS-DA can therefore be used for the effective in-line detection of adsorption endpoints.

Conclusions

This study compared four representative multivariate analysis methods, including MBSD, DMBA-TS, SIMCA, and PLS-DA, to determine the endpoints of HM adsorption processes. SIMCA and PLS-DA were most effective for endpoint determination. The results of this study provide a more comprehensive overview of the various multivariate analysis methods to facilitate the selection of the most suitable one. This study was conducive to addressing the issues of the in-line detection of adsorption endpoints to guide practical HM manufacturing.

Conflict of interest statement

The authors declare no conflict of interest.

Funding

This work was supported by the National Natural Science Foundation of China (82104383) and National S&T Major Project of China (2012ZX09101201-003).

Author contributions

Cheng Jiang was participated in the writing (original draft) of the paper, the performance of the research and data analysis. Haibin Qu was participated in conceptualization, research design, and writing (review & editing) of the paper.

Ethical approval of studies and informed consent

This article does not contain any research on human or animal subjects performed by any of the authors.

Acknowledgments

None.

References

[1] Bai Y, Ma J, Zhu W, et al. Highly selective separation and purification of chicoric acid from Echinacea purpurea by quality control methods in macroporous adsorption resin column chromatography. *J Sep Sci* 2019;42(5):1027–1036.
[2] Cren EC, Cardozo L, Silva EA, et al. Breakthrough curves for oleic acid removal from ethanolic solutions using a strong anion exchange resin. *Sep Purif Technol* 2009;69(1):1–6.
[3] Fayaz M, Zarifi MH, Abdolrazzaghi M, et al. A novel technique for determining the adsorption capacity and breakthrough time of adsorbents using a noncontact high-resolution microwave resonator sensor. *Environ Sci Technol* 2017;51(1):427–435.
[4] Wu YJ, Jin Y, Li YR, et al. NIR spectroscopy as a process analytical technology (PAT) tool for on-line and real-time monitoring of an extraction process. *Vib Spectrosc* 2012;58:109–118.
[5] Geskovski N, Stefkov G, Gigopulu O, et al. Mid-infrared spectroscopy as process analytical technology tool for estimation of THC and CBD content in Cannabis flowers and extracts. *Spectrochim Acta A Mol Biomol Spectrosc* 2021;251:119422.
[6] Wu SJ, Qiu P, Li P, et al. A near-infrared spectroscopy-based endpoint determination method for the blending process of Dahuang soda tablets. *J Zhejiang Univ Sci B* 2020;21(11):897–910.
[7] Tewari J, Strong R, Boulas P. At-line determination of pharmaceuticals small molecule's blending end point using chemometric modeling combined with Fourier transform near infrared spectroscopy. *Spectrochim Acta A Mol Biomol Spectrosc* 2017;173:886–891.
[8] Khorasani M, Amigo JM, Bertelsen P, et al. Detecting blending end-point using mean squares successive difference test and near-infrared spectroscopy. *J Pharm Sci* 2015;104(8):2541–2549.
[9] Igne B, de Juan A, Jaumot J, et al. Modeling strategies for pharmaceutical blend monitoring and end-point determination by near-infrared spectroscopy. *Int J Pharm* 2014;473(1–2):219–231.
[10] Yang NL, Cheng YY, Qu HB. An approach to purifying process analysis of Chinese herbal extracts using NIRS. *Acta Chim Sinica* 2003;61(5):742–747.
[11] Yan B, Qu H. Multivariate data analysis of UV spectra in monitoring elution and determining endpoint of chromatography using polyamide column. *J Sep Sci* 2013;36(7):1231–1237.
[12] Jiang C, Liu Y, Qu HB. Data fusion strategy based on near infrared spectra and ultraviolet spectra for simultaneous determination of ginsenosides and saccharides in Chinese herbal injection. *Anal Methods-Uk* 2013;5:4467–4475.

[13] Blanco M, Cueva-Mestanza R, Cruz J. Critical evaluation of methods for end-point determination in pharmaceutical blending processes. *Anal Methods-Uk* 2012;4(9):2694–2703.
[14] Wu Y, Jin Y, Ding H, et al. In-line monitoring of extraction process of scutellarein from Erigeron breviscapus (vant.) Hand-Mazz based on qualitative and quantitative uses of near-infrared spectroscopy. *Spectrochim Acta A Mol Biomol Spectrosc* 2011;79(5):934–939.
[15] Wu HQ, Khan MA. Quality-By-Design (QbD): an integrated approach for evaluation of powder blending process kinetics and determination of powder blending end-point. *J Pharm Sci-Us* 2009;98(8):2784–2798.
[16] Costa EB, Silva RC, Espejo-Roman JM, et al. Chemometric methods in antimalarial drug design from 1,2,4,5-tetraoxanes analogues. *SAR QSAR Environ Res* 2020;31(9):677–695.
[17] Rukundo IR, Danao MC. Identifying turmeric powder by source and metanil yellow adulteration levels using near-infrared spectra and PCA-SIMCA modeling. *J Food Prot* 2020;83(6):968–974.
[18] Puchert T, Holzhauer CV, Menezes JC, et al. A new PAT/QbD approach for the determination of blend homogeneity: combination of on-line NIRS analysis with PC scores distance analysis (PC-SDA). *Eur J Pharm Biopharm* 2011;78(1):173–182.
[19] Jimenez-Carvelo AM, Martin-Torres S, Ortega-Gavilan F, et al. PLS-DA *vs* sparse PLS-DA in food traceability. A case study: authentication of avocado samples. *Talanta* 2021;224:121904.
[20] Vieira LS, Assis C, de Queiroz M, et al. Building robust models for identification of adulteration in olive oil using FT-NIR, PLS-DA and variable selection. *Food Chem* 2021;345:128866.
[21] Jiang C, Qu H. A comparative study of using in-line near-infrared spectra, ultraviolet spectra and fused spectra to monitor Panax notoginseng adsorption process. *J Phar Biomed Anal* 2015;102:78–84.
[22] Jiang C, Gong XC, Qu HB. A strategy for adjusting macroporous resin column chromatographic process parameters based on raw material variation. *Sep Purif Technol* 2013;116:287–293.
[23] Jiang C, Gong X, Qu H. Multivariate modeling and prediction of breakthrough curves for herbal medicine adsorption on column chromatography: a case study. *Sep Sci Technol* 2015;50(7):1030–1037.
[24] Tranas R, Lovvik OM, Tomic O, et al. Lattice thermal conductivity of half-Heuslers with density functional theory and machine learning: enhancing predictivity by active sampling with principal component analysis. *Comp Mater Sci* 2022;202:110938.
[25] Wold S, Esbensen K, Geladi P. Principal component analysis. *Chemometr Intell Lab* 1987;2(1–3):37–52.
[26] Wold S. Pattern-recognition by means of disjoint principal components. *Pattern Recogn* 1976;8(3):127–139.
[27] Ledauphin J, Le Milbeau C, Barillier D, et al. Differences in the volatile compositions of French labeled brandies (Armagnac, Calvados, Cognac, and Mirabelle) using GC-MS and PLS-DA. *J Agr Food Chem* 2010;58(13):7782–7793.
[28] Frank IE, Friedman JH. A statistical view of some chemometrics regression tools. *Technometrics* 1993;35(2):109–135.
[29] Luis ML, Garcia JM, Jimenez F, et al. Simultaneous determination of chlorthalidone and spironolactone with univariate and multivariate calibration: wavelength range selection. *J AOAC Int* 1999;82(5):1054–1063.
[30] Liu W, Zhang S, Zu YG, et al. Preliminary enrichment and separation of genistein and apigenin from extracts of pigeon pea roots by macroporous resins. *Bioresource Technol* 2010;101(12):4667–4675.

How to cite this article: Jiang C, Qu HB. In-line spectroscopy combined with multivariate analysis methods for endpoint determination in column chromatographic adsorption processes for herbal medicine. Acupunct Herb Med 2022;2(4):253–260. doi: 10.1097/HM9.0000000000000035