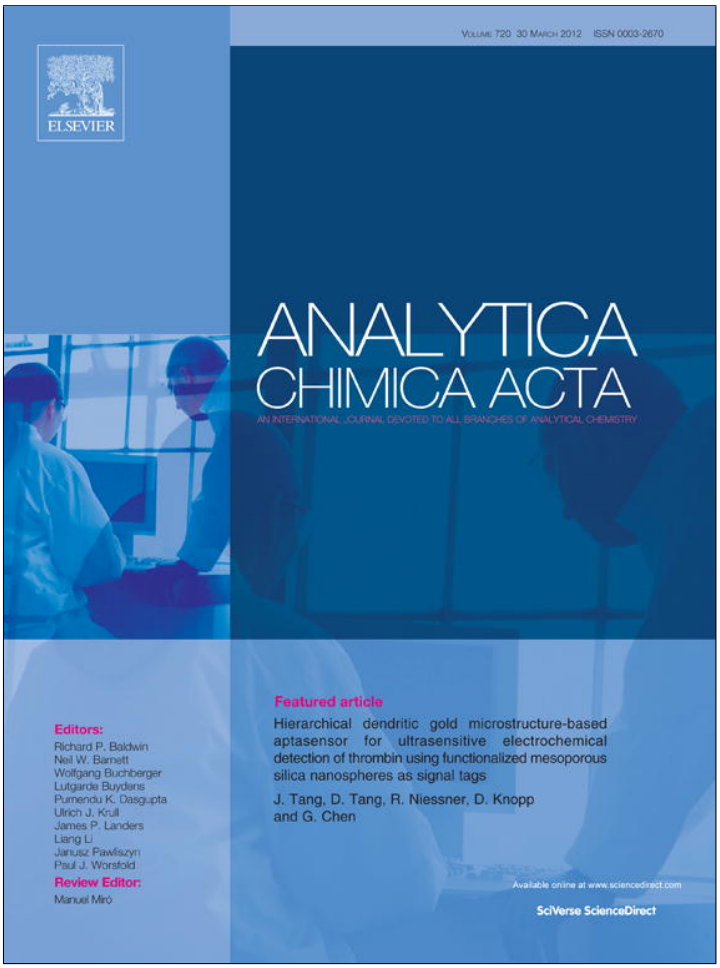

Analytica Chimica Acta 720 (2012) 22–28

Contents lists available at SciVerse ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

NIR analysis for batch process of ethanol precipitation coupled with a new calibration model updating strategy

Bing Xu[a,b], Zhisheng Wu[a,b], Zhaozhou Lin[a,b], Chenglin Sui[a,b], Xinyuan Shi[a,b,*], Yanjiang Qiao[a,b,*]

[a] Beijing University of Chinese Medicine, Beijing 100102, China
[b] The Key Laboratory of TCM-Information Engineering of State Administration of Traditional Chinese Medicine, Beijing 100102, China

ARTICLE INFO

Article history:
Received 16 November 2011
Received in revised form 8 January 2012
Accepted 13 January 2012
Available online 21 January 2012

Keywords:
Ethanol precipitation
Near infrared spectroscopy
Calibration model updating
Simple interval calculation
Flos Lonicerae Japonicae
Process analytical technology

ABSTRACT

Ethanol precipitation plays a major role in the pretreatment of Flos Lonicerae Japonicae of Qingkailing injection, and is also one of the most popular purification techniques in Chinese herbal medicines. In order to monitor and have a better understanding of the ethanol precipitation process, a PLS model was built based on NIR spectroscopy and HPLC analysis of chlorogenic acid content within the framework of FDA's PAT initiative. Nevertheless, due to the complex mechanism of and the raw materials' natural variability introduced into the ethanol precipitation process, it was unable to foresee the variations in new batches which may jeopardize the robustness of the established model. Therefore, based on the simple interval calculation (SIC) theory, a new model expansion updating strategy which could continuously expand the variation coverage of the calibration model along with the batch proceeding of ethanol precipitation process was proposed. Effects of model updating were validated by an individual batch with 60 samples. After two times of updating, the root mean squared error of prediction (RMSEP) decreased from 0.268 mg mL^{-1} to 0.199 mg mL^{-1}, while the insiders in the object status plot (OSP) increased from 44 to 58, demonstrating the good performance of the proposed approach.

1. Introduction

Ethanol precipitation is one of the most popular purification techniques used in the preparation process of traditional Chinese medicine (TCM). In Chinese Pharmacopeia (Ch.P. 2010 Edition, Volume I), about 140 manufacturing processes of Chinese patent medicine and simple preparations employ ethanol precipitation. Flos Lonicerae Japonicae is one of the ingredients in Qingkailing injection specified in Ch.P. and is widely used in the clinic. In the manufacturing process of Qingkailing injection, ethanol precipitation plays a major role in the pretreatment of Flos Lonicerae Japonicae, and is thus taken as our research object.

The principles of ethanol precipitation are easy to understand. During the addition of ethanol into the water extract of crude drugs, unwanted components like starch, inorganic acid salt, polysaccharide and proteins which are poorly dissolved in ethanol are precipitated, while the effective chemicals with good solubility in both water and ethanol are preserved. Nevertheless, the depositing particles could absorb the effective compositions due to the mechanism of electric effect and (or) hydrogen bonding interaction [1], leading to the loss of effective components which may affect the efficacy and safety of the yielded product. Besides, many process parameters could influence the ethanol precipitation, such as the temperature and density of water extract, speed of agitation and addition of ethanol, the final content of ethanol. Although these parameters can be operated well, the size and status of depositing particles could not be controlled so far, indicating that there exist some uncertainties in the ethanol precipitation process. What is more, the quality of raw materials could vary due to the different geographic origin and growing conditions [2]. Such natural variability may be introduced into the ethanol precipitation process, causing fluctuations between different batches.

Therefore, in order to guarantee the stable therapeutic effect and safety of the TCM preparations, NIR spectroscopy (NIRS) in line with the United States Food and Drug Administration (FDA)'s Process Analytical Technology (PAT) initiative and guidance [3], was applied in process monitoring of ethanol precipitation. NIRS is a fast and non-destructive technique that requires minimal analyst intervention [4], and recently has been more and more reported to be successfully applied in TCM research [5,6]. One objective of this study is to investigate the possibility of NIRS to quantify the content of chlorogenic acid, one of the most important effective components in both Flos Lonicerae Japonicae and Qingkailing injection [7], during the ethanol precipitation process of water extract of Flos Lonicerae Japonicae.

* Corresponding authors at: School of Chinese Pharmacy, Beijing University of Chinese Medicine, South of Wangjing Middle Ring Road, Chaoyang District, Beijing City 100102, China. Tel.: +86 010 84738621; fax: +86 010 84738661.
E-mail addresses: shixinyuan01@163.com (X. Shi), yjqiao@yahoo.com (Y. Qiao).

0003-2670/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.aca.2012.01.022

Firstly, multivariate calibration (MVC) tools (e.g. PLS) are used to build the quantitative model relating NIR spectra to the concentration of chlorogenic acid. Samples included in the initial calibration model should ideally encompass every source of variability that potentially influences measurement in the process [8], whereas it is difficult to foresee the new sources of variation in actual progress of ethanol precipitation due to the above-mentioned complex nature of the process. Such variation may jeopardize the model robustness over time. Hence, on second thoughts, a dynamic model updating strategy that continuously expands the variation coverage of the quantitative model along with the batch proceeding of ethanol precipitation process is proposed.

As reported, there are several calibration model maintenance strategies focusing on data correction and adjustment such as Slope and Bias Standardization (SBC) [9], Piecewise Direct Standardization (PDS) [10], Finite Impulse Response (FIR) [11], Spectral Space Transformation (SST) [12], Systematic Prediction Error Correction (SPEC) [13], model expansion, model modification (e.g. Dynamic Orthogonal Projection [14]) and new MVC algorithms (e.g. Delaunay Triangulation [15]). This work chiefly considers the model expansion technique. The simplest way of model expansion is adding new samples to the old calibration set to reconstruct a new model [16]. However, new samples incorporated into the old model should be identified with new sources of variation firstly. The Kennard and Stone algorithm, Duplex algorithm, Mahalanobis Distance (MD) [17] and Principal Component Analysis (PCA) [18] were reported to perform the choice of the new subsets. And this article, for the first time, couples the simple interval calculation (SIC) method [19] into the calibration model updating strategy. The SIC leverage is used to evaluate the extent of variability of incoming new samples. The criterion of SIC leverage for selecting updating samples is built.

2. Experimental

2.1. Materials

Flos Lonicerae Japonicae was purchased from Beijing Ben Cao Fang Yuan Medicine Co. Ltd. (Beijing, P.R. China). Chlorogenic acid was obtained from National Institute for the Control of Pharmaceutical and Biological Products (NICPBP, Beijing, P.R. China). Methanol (Fisher Scientific, USA), formic acid (Sigma–Aldrich Co. LLC, USA) and distilled water (A.S. Watson Group Ltd., Hong Kong of China) were of HPLC grade and all other reagents were of analytical grade.

2.2. Instrumentation

2.2.1. NIR spectroscopy

The NIR spectra were collected in transmission mode with an Antaris Nicolet FT-NIR system (Thermo Fisher Scientific Inc., USA). Each sample spectrum was the result of 16 scans in the range between 10,000 and 4000 cm^{-1} at ambient temperature using 4 cm^{-1} resolution, and was recorded as absorbance with air as reference. Every sample was scanned three times, and the final spectrum used for each sample was an average of the three. All NIR spectra were collected and archived using the Thermo Scientific Result software.

2.2.2. HPLC analysis

The concentration of chlorogenic acid was quantitated by RP-HPLC method recommended by the Ch.P. (2010 Edition, Volume I). An Agilent 1100 HPLC system (Agilent Technologies, USA) with a vacuum degasser, a quaternary pump, an autosampler, a thermostatic column compartment, a diode array detector (DAD) were used. Separation was performed on Agilent XDB C18 column (250 mm × 4.6 mm with 5 μm particle size) at 30 °C. The mobile phase were methanol and 0.1% formic acid water solution (20:80, v/v), and the signal was monitored at 327 nm. The flow rate was maintained at 1 mL min^{-1}.

The samples after NIR scanning were diluted by 50% methanol–water solution properly. Then the diluent was filtered through a Millipore membrane filter with an average pore diameter of 0.45 μm, and 10 μL filtrate was injected into the HPLC system for analysis.

2.3. Procedures

2.3.1. Process description

The Flos Lonicerae Japonicae was first extracted by 15 fold water for 0.5 h. The water extract was filtered. The remaining Flos Lonicerae Japonicae was extracted by 10 fold water for 0.5 h, and then filtered. Filtrates after the two extractions were combined and concentrated to certain densities (e.g. 1.10 at 25 °C for normal batches).

The ethanol precipitation process was carried out in a 3 L glass reactor. Reactor stirring was provided using an agitator with oblique quadricuspid paddle at a constant speed of 500 rpm. Ethanol was pumped into the reactor from the ethanol tank with a constant flow rate of 75 mL min^{-1}. In each batch of the ethanol precipitation process, the agitator and the pump were turned on simultaneously immediately after 400 mL concentrated water extract was added into the reactor. The agitation process was lasted for 30 min. During the process, when the predefined amount of ethanol (e.g. 1500 mL for normal batches) was acquired, the pump was turned off to stop the addition of ethanol. The final concentration of ethanol in normal batch was about 75%.

Along with the ethanol precipitation process, sample of 1.5 mL was drew by a pipette gun from the reactor at the fixed position every 30 s. Consequently, a total of 60 samples in one batch were prepared. All samples were centrifuged at 9000 rpm for 10 min. Then, the sample was filled into a particular quartz cuvette (8 mm in diameter) covered with a plastic cap to perform the NIR spectroscopy.

2.3.2. NIR spectra pretreatment

A variety of spectroscopic data preprocessing methods are compared to separate the useful information from noise, such as Savitzky–Golay (S–G) smoothing, first order derivatives, Norris Gap derivatives, multiplicative scatter correction (MSC), standard normal variate transformation (SNV), wavelet de-nosing of spectra (WDS). Then, moving window PLS (mwPLS) was applied in variable selection. SIMCA P + 11.5 (Umetrics, Sweden) served as chemometric tool for data preprocessing, and iToolbox [20] for variable selection.

The following parameters were calculated to evaluate the success of data pretreatment: correlation coefficient r for both calibration and validation sets, the root mean squared error of calibration (RMSEC), the root mean squared error of cross-validation (RMSECV), the root mean squared error of prediction (RMSEP) [21] and residual predictive deviation (RPD) [22].

2.3.3. Model building

Six batches of ethanol precipitation process were performed. The former three batches (Batches 1, 2 and 3) were applied to form the initial calibration set. Batch 4 was applied as the internal test set. Batches 5 and 6 were used as the external test sets. In order to construct a calibration set covering a wide range of variations, water extracts for Batches 1 and 2 were concentrated to densities of 1.05 and 1.15 respectively. The amount of ethanol added to Batches 1 and 2 were both 1711 mL to make final concentration of ethanol in the two batches around 77%. Batches 3 to 6 were normal batches.

Table 1
Comparison of different preprocessing methods for NIR spectra.

Processing	LVs	Calibration				Validation			
		r_{cal}	RMSEC	RMSECV	$BIAS_{cal}$	r_{val}	RMSEP	RPD	$BIAS_{val}$
Origin	4	0.8997	0.606	0.640	0.482	0.9968	0.228	5.80	0.168
S-G	7	0.9287	0.515	0.613	0.398	0.9871	0.297	4.46	0.251
1std	6	0.8808	0.658	0.730	0.517	0.9937	0.343	3.85	0.290
NG-2ndd	6	0.8718	0.684	0.889	0.499	0.9117	0.656	2.01	0.500
MSC	3	0.8859	0.644	0.672	0.507	0.9967	0.300	4.41	0.228
SNV	4	0.8884	0.637	0.672	0.509	0.9973	0.285	4.64	0.225
WDS	4	0.8997	0.606	0.640	0.482	0.9969	0.228	5.80	0.168
Baseline	4	0.8884	0.638	0.673	0.506	0.9963	0.295	4.48	0.245

Origin means using the original spectra; 1std is the first order derivative; NG-2ndd is Norris Gap smoothing plus the second derivatives.

PLS regression performed on Matlab version 7.0 (MathWorks Inc., USA) with PLS_Toolbox 2.1 (Eigenvector Research Inc., USA) was used to set up quantitative models. The number of latent variables was optimized by leave-one-out (LOO) cross-validation method and predicted residual error sum square (PRESS). The performance of the built PLS model was evaluated in terms of r, RMSEC, RMSECV, RMSEP, BIAS and RPD.

2.3.4. SIC approach

Unlike the traditional MVC methods (e.g. PLS and PCR) whose assumption was based on error normality, the simple interval calculation (SIC) method was based on a single postulate that all errors involved in calibration problem were finite. The whole theory of SIC approach could be found in Ref. [23] and some applications in Refs. [24,25].

Consider the PLS regression as:

$$\mathbf{y} = \mathbf{Xb}+\boldsymbol{\varepsilon} \tag{1}$$

where $\mathbf{y}$ was the concentration of chlorogenic acid; $\mathbf{X}$ was a matrix of spectra with the dimensionality $n \times p$; n denote the number of samples while p the number of wavelength variables; $\mathbf{b}$ was the regression coefficients and $\boldsymbol{\varepsilon}$ was the prediction error.

According to the SIC theory, there was a maximum error deviation (MED, represented by β) of $\boldsymbol{\varepsilon}$. Thus, for a given calibration set ($\mathbf{X}$, $\mathbf{y}$), the following equation could be established:

$$\mathbf{B} = \{\mathbf{b} \in \mathbf{R}^p, \mathbf{y}^- < \mathbf{Xb} < \mathbf{y}^+\},$$

$$\text{where } y_i^- = y_i - \beta, y_i^+ = y_i + \beta, i = 1, \ldots, n \tag{2}$$

$\mathbf{B}$ in the above inequation was called the region of possible values (RPV) that stood for the parameter's space. When the spectrum ($\mathbf{x}$) of a new sample was acquired, the predicted concentration of chlorogenic acid $\mathbf{y} = \mathbf{x}^t\mathbf{b}$ would fall into the interval:

$$v = [v^-, v^+], \quad \text{where } v^- = \min(\mathbf{x}^t\mathbf{b}), v^+ = \max(\mathbf{x}^t\mathbf{b}) \tag{3}$$

The key to estimate the interval v was the calculation of β [23]. In this work, only the upper limit of β (β_{max}) was utilized. β_{max} approximates 4 times of RMSEC, giving an assurance that any samples will never cross outside this border [19]. The interval v and β_{max} were estimated using the SIC Toolbox [26].

The quality of SIC prediction was evaluated by two statistics. The SIC residual (r) was defined as the difference between the center of the interval v and the reference value y (scaled by β_{max}).

$$r(x, y) = \frac{1}{\beta}\left(y - \frac{v^+(x) + v^-(x)}{2}\right) \tag{4}$$

The SIC leverage (h) was the half width of interval v divided by β_{max}.

$$h(x, y) = \frac{1}{\beta}\left(\frac{v^+(x) - v^-(x)}{2}\right) \tag{5}$$

Based on the two statistics, an object status classification (OSC) [27] strategy was brought forward for both calibration set and test set. According to OSC, samples with different properties could be classified into insiders ($|r| + |h| < 1$), boundary objects ($|r| + |h| = 1$), outsiders ($|r| + |h| > 1$), absolute outsiders ($|h| > 1$) and outliers ($|r| - |h| > 1$).

As seen in Eq. (5), the SIC leverage characterized the β normalized precision of prediction, samples with larger SIC leverage meant that the prediction was poor and indicated the existence of variations. So, the calibration model, in this work, was updated by expanding with new samples selected according to the SIC leverage. The efficiency of updating was compared with an updating strategy using Kennard and Stone algorithm.

3. Results and discussion

3.1. HPLC determination of chlorogenic acid

For quantitative consideration, a calibration curve based on the concentration range from 2.01 to 80.40 $\mu g\,mL^{-1}$ of chlorogenic acid was established upon eight consecutive injections of different concentrations. Regression equation calibrated was $y = 16.4410x - 2.3636$ ($r = 0.9998$, $n = 8$) with y being the peak area in mAU s and x being the concentration. Quantification results showed that the range of chlorogenic acid concentrations of calibration set varied from 0.9319 $mg\,mL^{-1}$ to 7.7465 $mg\,mL^{-1}$. For test sets 4, 5 and 6, the ranges were 1.4103–6.4637 $mg\,mL^{-1}$, 1.4946–6.5796 $mg\,mL^{-1}$ and 1.4819–6.5266 $mg\,mL^{-1}$, respectively. It was seen that the initial calibration set covered a wide concentration range, within which concentrations of samples in the three test sets spread.

3.2. Data pretreatment

As seen in Table 1, the original NIR spectra (see Fig. 1) without any preprocessing were found to bear the good capability in both calibration and validation. In mwPLS, the spectra window size was set at 21 and the RMSECV for each window was calculated by PLS algorithm (see Fig. 2). The wavelength region of 8500–7300 cm^{-1} with relative low RMSECV was selected and applied in the following study.

3.3. Calibration and validation of quantitative model

The calibration results (see Fig. 3) indicated that six latent factors with the cumulative PRESS of 10.307 were enough. The validation results of Batches 4, 5 and 6 are shown in Table 2. It could be found that the RMSEP and BIAS of Batches 5 and 6 were larger than Batch 4, while the RPD decreased. The phenomenon of residual drift was observed in the residual plots of Batches 5 and 6 (see Fig. 4). The inferior prediction capability of the established model for new batches could be accounted for the uncertain variations between different batches. On the other hand, the result demonstrated a

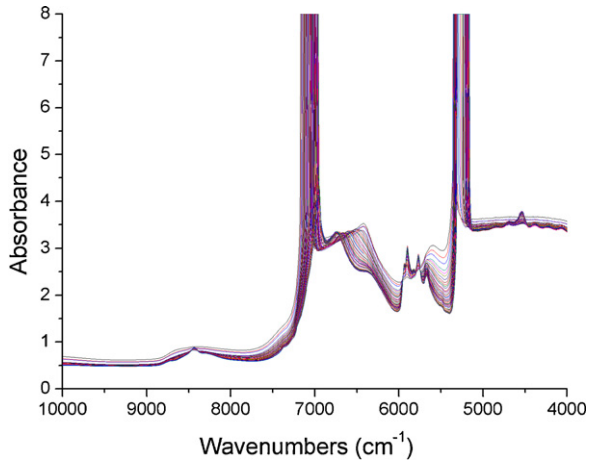

Fig. 1. Raw NIR spectra of Batch 3.

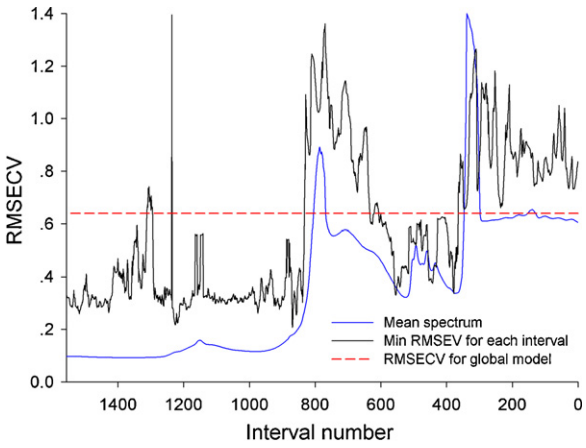

Fig. 2. Moving window PLS regression. Red dash line is RMSECV (4 latent factors) for the global model. Blue line is the mean spectrum scaled by the vertical axis. Black line represents the minimum RMSECV for each interval. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Validation statistics of Batches 4, 5 and 6.

Batch	r_{val}	RMSEP	RPD	$BIAS_{val}$
4	0.9940	0.156	8.45	0.111
5	0.9962	0.377	3.61	0.313
6	0.9944	0.320	4.04	0.279

calibration model updating along with the proceeding of the production process to improve the performance of the model.

3.4. Development of calibration subset

Since the initial model built may contain some useless and redundant spectral data or samples, a representative calibration subset was constructed by boundary samples in OSC before model updating. A theoretical discussion on the advantages and rationality of the representative subset selection in such a way could be referred to Ref. [27]. According to Eqs. (4) and (5), the SIC residuals and SIC leverages were computed. Then 51 boundary objects (green circles in Fig. 5) out of 180 samples in the initial calibration set were designated as the most significant objects for modeling, and the remaining samples (black squares in Fig. 5) may be deemed as redundant.

The predictability of the new PLS model constructed by the 51 boundary samples using 6 latent factors was validated by test set from Batch 4. The obtained results were with RMSEP = 0.134, $BIAS_{val}$ = 0.101 and RPD = 9.89, revealing that the selected calibration subset could predict samples with a better accuracy and was of maximal representation with respect to the initial model.

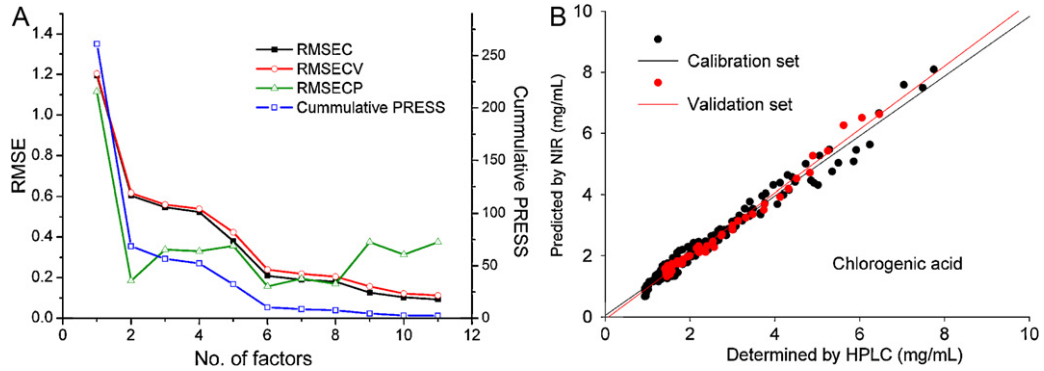

Fig. 3. (A) Calibration characteristics vs. number of latent factors; (B) correlation diagram for chlorogenic acid (validation set is from Batch 4).

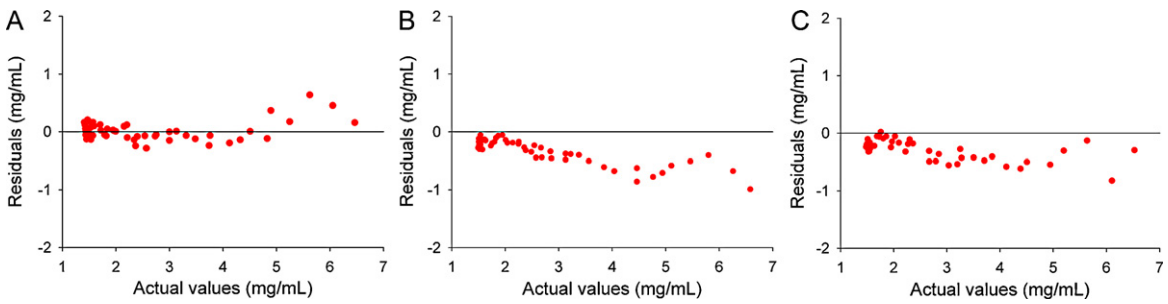

Fig. 4. (A) Residuals plot of Batch 4; (B) residuals plot of Batch 5; (C) residuals plot of Batch 6.

Table 3
Comparison of the performance of the updated models validated by Batch 6.

Strategy	Model	No. of samples in calibration set	LVs	Calibration			Validation		
				RMSEC	RMSECV	$BIAS_{cal}$	RMSEP	RPD	$BIAS_{val}$
1	A	51	6	0.330	0.399	0.263	0.268	4.82	0.224
	B	111	6	0.249	0.279	0.175	0.261	4.95	0.221
	C	163	6	0.255	0.272	0.188	0.199	6.49	0.174
2	I	180	6	0.209	0.239	0.157	0.320	4.04	0.279
	II	200	6	0.211	0.230	0.159	0.314	4.12	0.277
	III	220	6	0.233	0.246	0.183	0.283	4.57	0.250

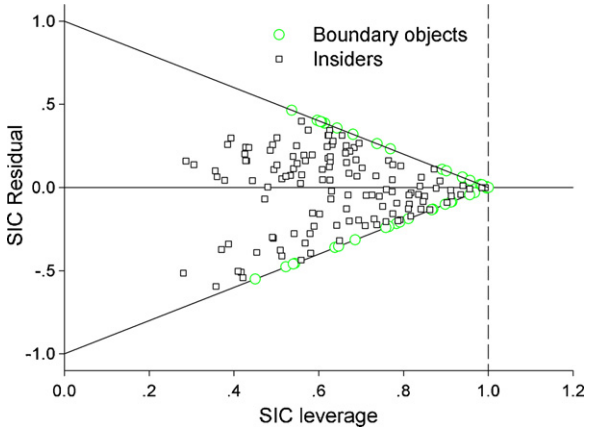

Fig. 5. SIC object status plot for the initial calibration set. (For interpretation of the references to colour in the text, the reader is referred to the web version of this article.)

3.5. Model updating

Now, it is assumed that only the initial quantitative model (represented by Model I consisting of 180 samples) and its representative subset (represented by Model A consisting of 51 boundary samples) were accomplished. Then, two strategies of model updating were considered as follows:

(1) when a new batch (e.g. Batch 4) was performed, the SIC leverage could be computed through the NIR spectra. The samples in Batch 4 with SIC leverage larger than 0.5 were selected out and were quantified by HPLC. Afterwards, the spectra data and the corresponding reference values of the selected samples were added into Model A to rebuild a new calibration model named as Model B. Similarly, Model C was updated by adding the preferred samples from Batch 5. According to Eq. (5), the criteria of SIC leverage 0.5 was set to incorporate the samples with SIC interval larger than a half of β_{max} that approximated two times of RMSEC.

(2) The Kennard and Stone algorithm was employed to perform the subset selection. In our case, 20 samples were selected by the KS algorithm from a new batch and were added into the old model to create the updated one.

It could be imagined that both updating methods could be performed along with the batch proceeding of the ethanol precipitation process, resulting in two model series. In this work, the primary model I and Model A were updated twice. Number of samples in calibration set of the two series of models could be read in Table 3. Batch 6 was employed to validate and analyze the performance of these updated models. The calibration and validation statistics are also shown in Table 3, from which it was seen that in each model series, the updated model performed better than the former one with the increased RPD and decreased RMSEC, RMSECV, RMSEP and BIAS. Moreover, models updated by strategy (1) with less calibration samples exerted better predictability than that by strategy (2).

A further investigation was carried out by SIC analysis (see Table 4). The object status plots (OSP) (see Fig. 6) were drew to visually display the effectiveness of the updated models. Results revealed that more and more samples were classified as insiders. After two times of updating, there were no absolute outsiders. In view of the overall situation, samples seemed to be moving toward the Y axis during the course of updating, which was proved by the decreasing mean values of SIC leverage (h_{mean}) in each model series in Table 4. And the mean width of prediction interval (w_{mean}) in each series becomes smaller and smaller, demonstrating that uncertainty was reduced while the model was updated. β_{max} showed a trend of decreasing in Model series A–C, which agreed with a general concept that modeling error should reduce when model is expanded. However, such tendency was a little obscure in Model series I–III. After all, updating strategy (1) exhibited better performance with relatively smaller w_{mean} and h_{mean} as well as more insiders classified than strategy (2).

In addition, numbers of samples updated in our case by strategy (1) for the first and second time were 60 and 52, respectively. Nevertheless, if the model updating continued, the SIC leverages of the next Batch 6 could be calculated based on Model C. And it would be found that only 9 samples were with SIC leverages larger than

Table 4
General characteristics of the updated models by SIC analysis.

Strategy	Model	β_{max}	β_{max}/RMSEC	w_{mean}	h_{mean}	No. of insiders	No. of outsiders	No. of absolute outsiders	No. of boundary objects
1	A	1.0197	4.87	1.2217	0.5991	45	13	2	51
	B	1.1112	5.27	1.3495	0.6072	46	12	2	59
	C	0.8512	3.65	0.7474	0.4391	53	7	0	41
2	I	1.2241	3.71	1.908	0.7793	44	13	3	49
	II	1.1452	4.6	1.5244	0.6699	46	12	2	45
	III	0.8717	3.41	0.7292	0.4183	58	2	0	34

w_{mean} and h_{mean} represent the mean value of SIC intervals and SIC leverages for Batch 6, respectively. No. of insiders, outsiders and absolute outsiders regard to the OSC results of Batch 6 based on the corresponding model. No. of boundary objects regards to the OSC result of the calibration set.

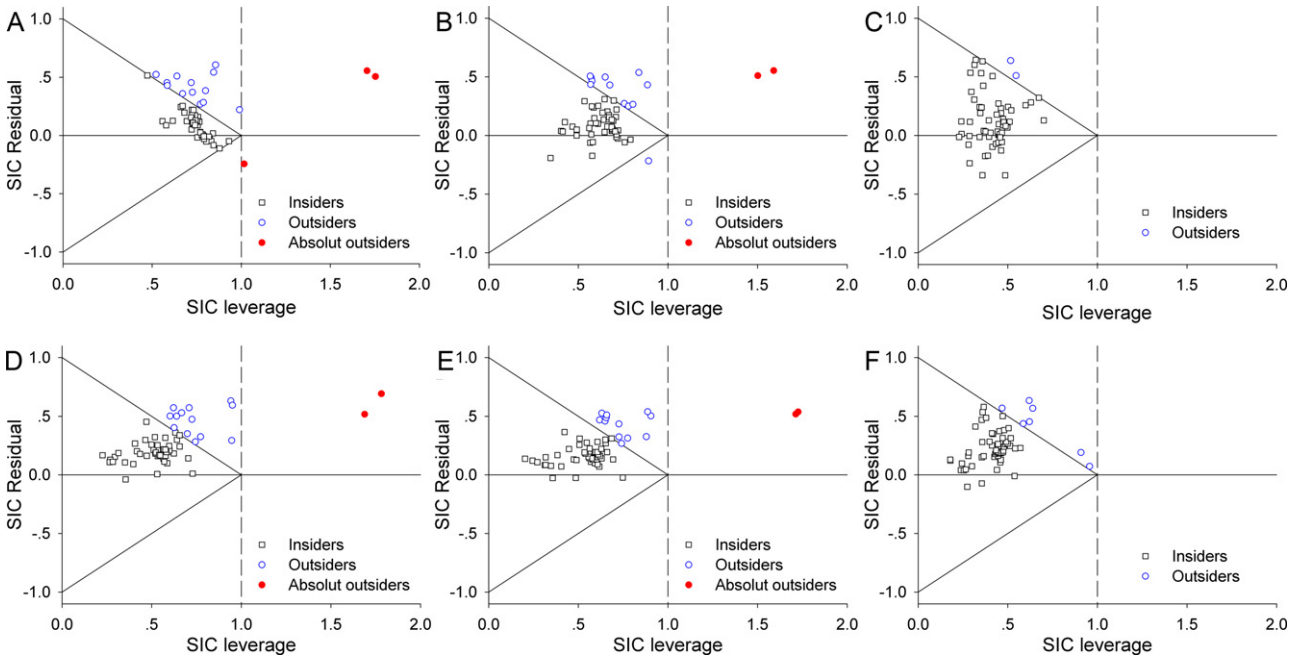

Fig. 6. SIC object status plot of Batch 6 (A) based on Model A; (B) based on Model B; (C) based on Model C; (D) based on Model I; (E) based on Model II; (F) based on Model III.

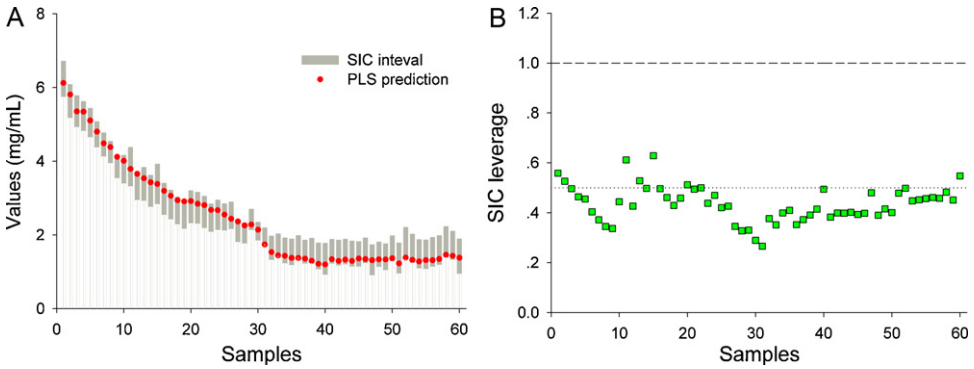

Fig. 7. Prediction quality for Batch 7 based on Model D with $\beta_{max} = 0.8711$: (A) PLS prediction (closed red dots) and SIC prediction intervals (blue bars) for samples of Batch 7; (B) calculated SIC leverages of samples of Batch 7. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the predefined criteria of 0.5. Then Model D was constructed consisting of 172 objects, the prediction capability of which was tested by Batch 7 without reference analysis. The prediction quality for Batch 7 based on Model D could be viewed in Fig. 7. The result was favorable, indicating the efficiency and robustness of the strategy (1) suggested.

4. Conclusion

Based on the PAT concept and NIR technology, a PLS model relating NIR spectra to HPLC references was successfully built to determine the low concentration of chlorogenic acid during the ethanol precipitation process of water extract of Flos Lonicerae Japonicae. However, when the developed quantitative model was validated by new batches, a phenomenon of residual drift was observed due to the batch-to-batch variations. In order to cope with the fluctuations between different batches, a new model updating strategy based on SIC theory was presented. Different from the existing model expansion methods who select samples needed to be updated randomly or only by the information of NIR data, the new strategy makes full utilization of the NIR spectra from the calibration and validation sets as well as the reference analysis results of the calibration set. Effects of model updating were evaluated by both the conventional statistics and SIC analysis. Results revealed that the proposed strategy outperformed the approach using Kennard and Stone algorithm. The course of model updating was visualized using object status plots (OSP), which would be helpful in applications.

Acknowledgements

The authors wish to acknowledge research funding supports from "the National Major Projects of Science and Technology named 'Creation of Major New Drugs' (No. 2010ZX09502-002, China)" and Scientific Research Project of Beijing University of Chinese Medicine (No. JYB22-XS034).

References

[1] Y. Chen, Y.C. Jin, S. Cai, Y.Y. Cheng, H.B. Qu, World Sci. Technol. 9 (2007) 16–19.
[2] S.S. Rosa, P.A. Barata, J.M. Martins, J.C. Menezes, J. Pharm. Biomed. Anal. 47 (2008) 320–327.
[3] U.S. Food and Drug Administration, Guidance for Industry PAT: A Framework for Innovative Pharmaceutical Development, Manufacturing and Quality Assurance, 2004, http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM070305.pdf.

[4] T. De Beer, A. Burggraeve, M. Fonteyne, L. Saerens, J.P. Remon, C. Vervaet, Int. J. Pharm. 417 (2011) 3–47.
[5] Y. Jiang, B. David, P.F. Tu, Y. Barbin, Anal. Chim. Acta 657 (2010) 9–18.
[6] Y.W. Wu, S.Q. Sun, Q. Zhou, H.W. Leung, J. Pharm. Biomed. Anal. 46 (2008) 498–504.
[7] H.Y. Zhang, P. Hu, G.A. Luo, Q.L. Liang, Y.L. Wang, Y.M. Wang, Anal. Chim. Acta 577 (2006) 190–200.
[8] M. Blanco, M.A. Romero, M. Alcalà, Talanta 64 (2004) 597–602.
[9] Y.D. Wang, D.J. Veltkamp, B.R. Kowalski, Anal. Chem. 63 (1991) 2750–2756.
[10] E. Bouveresse, C. Hartmann, D.L. Massart, Anal. Chem. 68 (1996) 982–990.
[11] T.B. Blank, S.T. Sum, S.D. Brown, Anal. Chem. 68 (1996) 2978–2995.
[12] W. Du, Z.P. Chen, L.J. Zhong, S.X. Wang, R.Q. Yu, A. Nordon, D. Littlejohn, M. Holden, Anal. Chim. Acta 690 (2011) 64–70.
[13] Z.P. Chen, L.M. Li, R.Q. Yu, D. Littlejohn, A. Nordon, J. Morris, A.S. Dann, P.A. Jeffkins, M.D. Richardson, S.L. Stimpson, Analyst 136 (2011) 98–106.
[14] M. Zeaiter, J.M. Roger, V. Bellon-Maurel, Chemom. Intell. Lab. Syst. 80 (2006) 227–235.
[15] L. Jin, Q.S. Xu, J. Smeyers-Verbeke, D.L. Massart, Chemom. Intell. Lab. Syst. 80 (2006) 87–98.
[16] W.L. Li, L.H. Xing, L.M. Fang, J. Wang, H.B. Qu, J. Pharm. Biomed. Anal. 53 (2010) 350–358.
[17] X. Capron, B. Walczak, O.E. de Noord, D.L. Massart, Chemom. Intell. Lab. Syst. 76 (2005) 205–214.
[18] M. Blanco, R. Cueva-Mestanza, A. Peguero, Talanta 85 (2011) 2218–2225.
[19] O.Y. Rodionova, A.L. Pomerantsev, Chemom. Intell. Lab. Syst. 97 (2009) 64–74.
[20] L. Nørgaaard, iToolbox for MATLAB, 2005, http://www.models.life.ku.dk/iToolbox.
[21] M.C. Sarraguça, J.A. Lopes, Vib. Spectrosc. 49 (2009) 204–210.
[22] C. Collell, P. Gou, J. Arnau, J. Comaposada Food Chem. 129 (2011) 601–607.
[23] A.L. Pomerantsev, Progress in Chemometrics Research, Nova Science Publisher, New York, 2005, pp. 43–64.
[24] A. Pomerantsev, O. Rodionova, A. Höskuldsson, Chemom. Intell. Lab. Syst. 81 (2006) 165–179.
[25] A.L. Pomerantsev, O. Ye Rodionova, Chemom. Intell. Lab. Syst. 79 (2005) 73–83.
[26] Semenov Institute of Chemical Physics, Software Implementation of SIC Method for MATLAB, 2008, http://rcs.chph.ras.ru/sic.
[27] O.Y. Rodionova, K. Esbensen, A. Pomerantsev, J. Chemom. 18 (2004) 402–413.