Talanta 107 (2013) 248–254

Contents lists available at SciVerse ScienceDirect

Talanta

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

A novel model selection strategy using total error concept

Zhisheng Wu [a,b,c], Qun Ma [a,b,c], Zhaozhou Lin [a,b,c], Yanfang Peng [a,b,c], Lu Ai [a,b,c], Xinyuan Shi [a,b,c,*], Yanjiang Qiao [a,b,c,*]

[a] Beijing University of Chinese Medicine, Beijing 100102, China
[b] Key Laboratory of TCM-information Engineering of State Administration of TCM, Beijing 100102, China
[c] Beijing Key Laboratory for Basic and Development Research on Chinese Medicine, Beijing 100102, China

ARTICLE INFO

Article history:
Received 7 September 2012
Received in revised form
29 December 2012
Accepted 31 December 2012
Available online 10 January 2013

Keywords:
Near-infrared spectroscopy
Accuracy profile
Model selection
Partial least squares regression

ABSTRACT

Our previous work had proved that accuracy profile theory could be employed as a means of validating one PLS model in Chinese material medica system. In this paper, accuracy profile theory is proposed as a powerful decision tool to demonstrate the prediction performance of multi-model at each concentration level rather than all concentration levels. Partial least square (PLS), interval partial least square (iPLS), backward interval partial least square (BiPLS) and moving window partial least square (MWPLS) were selected to construct visible and near-infrared (vis/NIR) spectroscopy models. Chemometric indicators, i.e., determination coefficient (R^2), root-mean-square error of prediction (RMSEP) and ratio of performance to inter-quartile (RPIQ), were used to select the optimum model. However, the results clarified that these commonly used indicators could not clearly demonstrate different PLS models' ability because these indicators depend on all concentration levels to assess the multi-model. Therefore, "total error concept" (accuracy profile theory) was introduced to assess the ability of multi-model at each concentration level. Analytical methodology parameters, i.e., linearity, relative bias, uncertainty, repeatability, intermediate precision, lower limit of quantification (LLOQ) and risk, were calculated by accuracy profile theory. Final results showed that model selection strategy which was based on model assessment at every concentration level was more sensitive than the one based on all concentration levels. The analytical procedures involved in this work ensure that model selection strategy using total error concept is coherent.

1. Introduction

Near-infrared (NIR) spectroscopy has gained great attention to date as a Process Analytical Technology (PAT) tool because it can provide rapid information collection with minimal or no sample preparation, and has been widely used in the quantitative and qualitative analysis of pharmaceutical products [1–5]. Compared with mid-IR spectra, whose absorbance bands are directly interpretable due to chemical peak specificity, NIR spectra are difficult to interpret due to their nature (overtones and combination bands of vibrational energy levels). Consequently, a calibration model is required for NIR analysis [6].

The NIR linear model is mainly established based on partial least square (PLS). The theory and application of PLS in spectrometry have been discussed by several researchers [7]. Recently, a PLS model has been developed including interval partial least square (iPLS) [8], genetic algorithm-interval partial least square (GA-iPLS) [9], forward interval partial least square (FiPLS) [10], backward interval partial least-square (BiPLS) [10], genetic algorithm-partial least square (GA-PLS) [11], and moving window-partial least square (MWPLS) [12].

These different PLS algorithms were previously assessed according to the usual chemometric indicators: root mean square error (RMSE), standard error of prediction (SEP), ratio of performance to deviation (RPD), ratio of performance to inter-quartile (RPIQ) [13], determination coefficient (R^2) and bias. However, it has not been reported whether these indicators can be used to accurately assess PLS algorithm. In this paper, we clarified it, and presented a total error (systematic and random errors) concept to assess different PLS algorithms. Based on "β-expectation tolerance intervals", accuracy profile theory was introduced by considering the total error [14–17].

Besides, accuracy profile theory fully complies with the ICH Q2(R1) regulatory documents as it integrates all required analytical methodology parameters, i.e., linearity, relative bias, uncertainty, repeatability and intermediate precision, the lower limit of quantification (LLOQ) and risk [18–20]. Therefore, a novel model selection strategy was given using total error concept (accuracy profile). In this research, a multi-component matrix, Yinhuang oral solution was used as an example. Different PLS algorithms were assessed by chemometric indicators and accuracy profile.

* Corresponding authors at: Beijing University of Chinese Medicine, Beijing 100102, China. Tel.: +86 10 84738621; fax: +86 10 84738661.
E-mail addresses: shixinyuan01@163.com (X. Shi), yjqiao@263.net (Y. Qiao).

0039-9140/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.talanta.2012.12.057

2. Materials and methods

2.1. Materials

Yinhuang oral solution was purchased from Jvrong Pharmaceutical Group Co., Ltd. (Jiangsu, China) and deposited in the Key Laboratory of TCM-information Engineering of State Administration of Traditional Chinese Medicine (No. 110201). Baicalin reference standard (lot number: 110777–201005) was supplied by the National Institute for the Control of Pharmaceutical and Biological Products (Beijing, China). HPLC grade methanol was purchased from Tedia (USA). Deionized water was purified by a Milli-Q water system (Millipore Corp., Bedford, MA, USA).

2.2. Design of experiment

135 samples of three lots Yinhuang oral solution were obtained containing a series of baicalin concentration levels. An experimental protocol was created sequentially for the calibration and validation steps in order to obtain a robust model (Table 1). Seventy-two sample sets were obtained in the calibration set. The validation set (63 samples) was established in the same way.

2.3. NIR equipment and software

The visible and NIR spectra were collected in the transmission mode using an XDS rapid liquid analyzer with VISION software (Foss NIR Systems, Silver Spring, MD, USA). The sample was held in a circular sample cuvette with plastic cap (optical path is 8 mm). Each spectrum was the average of 32 scans with a wavelength increment of 0.5 nm. The range of spectra was from 400 nm to 2500 nm. Each sample was analyzed three times and the mean value of three spectra was used in the following analysis.

Data analysis was performed using home-made routines programmed in MATLAB code (MATLAB, The MathWorks, Massachussetts). The toolbox used to select the most informative variables, called iPLS, BiPLS and MWPLS, was downloaded from http://www.models.kvl.dk/. The calculation of the accuracy profile was adopted by e. noval V3.0 software (Arlenda, Liège, Belgium).

2.4. HPLC method

Amounts of baicalin were accurately weighed using an XS205DU electronic balance (Mettler Toledo, Greifensee, Schweiz), and dissolved with 5 mL of methanol. The reference method used for baicalin determination was the HPLC assay recommended by the Chinese Pharmacopoeia (ChP, 2010 Edition) for Yinhuang oral solution. An Aglient 1100 series HPLC apparatus, equipped with a quaternary solvent delivery system, an autosampler and a DAD detector, was used. The concentration of baicalin was analyzed by reverse-phase chromatography on an ODS column (150 mm × 4.6 mm, 5 μm, Waters, USA) with isocratic elution of the mobile phase consisting of methanol, water and phosphoric acid (50:50:0.2, v/v) at a flow rate of 1.0 mL/min. A column temperature of 30 °C, and detection wavelength at 274 nm were set.

Table 1
Variability sources included in the calibration and validation sets.

	Calibration set	Amount of variability	Validation set
Variability sources			
Lots		3	
Operators	2		2
Days	2		2
NIR acquisitions	3		3

2.5. Theory and algorithm

2.5.1. iPLS model

The iPLS algorithm used was downloaded from the Royal Veterinary and Agricultural University of Denmark. The principle of iPLS algorithm is to split the spectra into smaller equidistant regions and, afterward, to develop PLS regression models for each of the sub-intervals, using the same number of latent factors.

Thereafter, an average error is calculated for every sub-interval and for the full-spectrum model. The region with the lowest error is chosen. An optimized region can be found by subtracting or adding new factors. One of the main advantages of this method is the possibility to represent a local regression model in a graphical display, focusing on a choice of better intervals and permitting a comparison among interval models and the full-spectrum model. This method is intended to give an overview of the data and can be helpful in interpretation.

2.5.2. BiPLS model

The BiPLS algorithm used here was developed by Norgaard et al. [8]. As in the iPLS model, the data set is split into a given number of intervals, but now PLS models are calculated with each interval left out, i.e., if one chooses 40 intervals then each model is based on 39 intervals leaving out one interval each time. The first left out interval is the one that gives the poorest performing model with respect to root mean square error (RMSE). This procedure is continued until one interval remains.

2.5.3. MWPLS model

MWPLS is a modeling technique that can be thought of as a series of diagnostic PLS regressions based on all continuous window size "H" in the parent data set. In effect, a window of size H is "moved" across the data set to collect modeling information. The model quality and number of latent variables (LVs) required for model production during this process can then be used to find the best spectral region(s) of size H. The original MWPLS work uses the sum of squared residues (SSR) as a measure of a model fit:

$$SSR_j = \sum_{i=1}^{I} \left(Y_{i,j,pre} - Y_{i,j,ref}\right)^2 \quad (1)$$

where $Y_{i,j,pre}$ and $Y_{i,j,ref}$ are the reference and predicted values of sample i, respectively for PLS models constructed with i total samples and j LVs. This value is calculated based on PLS models of various LV constituency and window size H, with cross-validation calculations performed only on the most promising combinations of range and number of LVs.

2.5.4. Accuracy profile theory

The concept of accuracy profile is introduced based on the fitness-for-purpose approach of the validation. The basic idea is the acceptability limit criterion, noted by λ. It is assumed that end-users actually expect an analytical procedure to return a result $\hat{Z}$ which differs from the unknown target value Z by less than λ. This requirement can be expressed as

$$|Z - \hat{Z}| < \lambda \quad (2)$$

A procedure can be validated if it is very likely that the requirement given by Eq. (2) is fulfilled, i.e.

$$P\left(|Z - \hat{Z}| < \lambda\right) \geq \beta \quad (3)$$

β being the probability that a future determination falls inside the acceptability limits. It is possible to compute the so-called "β-expectation tolerance interval" (β-ETI). The β-ETI is given by

$$\delta \pm Q_t k_S S_R \tag{4}$$

Q_t is the β quantile of the Student's t-distribution, δ is the bias and k_s is the expansion factor:

$$k_S = \sqrt{1+\frac{1}{pnB^2}} \tag{5}$$

with

$$B = \sqrt{\frac{R+1}{nR+1}};\ R = \frac{S_B^2}{S_r^2} \tag{6}$$

S_B^2 and S_r^2 are the estimates of the between-conditions variance and the within-conditions variance (repeatability). The reproducibility variance, S_R^2 is obtained through

$$S_R^2 = S_r^2 + S_B^2 \tag{7}$$

It can be easily demonstrated that

$$k_S^2 S_R^2 = S_R^2 + \frac{nS_B^2 + S_r^2}{np} \tag{8}$$

The second member of the right term of Eq. (8) is the estimation of the variance of the overall mean that can be assimilated to the bias uncertainty for a nested design with p conditions for experiments and n replications within each condition. Thus

$$k_S^2 S_R^2 = S_R^2 + S_\delta^2 = u^2(Z) \tag{9}$$

and β-ETI can be given now by

$$\delta \pm Q_t u(Z) \tag{10}$$

Therefore, based on β-expectation tolerance intervals, the accuracy profile theory makes a visual and reliable representation for actual and future model performance possible.

2.5.5. Chemometric indicators

To express the advantages of NIR-calibration models, the optimum model was established based on RMSE, standard error of cross validation (SECV, five segment size), SEP, RPD (RPD=SD/SEP), and RPIQ et al. The RPIQ parameter was proposed based on quartiles by Veronique Bellon-Maurel, which better represents the spread of the population. The quartiles are milestones in the population range: Q1 is the value below which we can find 25% of the samples; Q3 is the value below which we find 75% of the samples; and, Q2, commonly called the median, is the value under which 50% of samples are found. The quartiles are therefore useful to determine equivalent ranges of population spread. For example, inter-quartile distance IQ ($Q3$–$Q1$) gives the range that accounts for 50% of the population around the median. The *RPIQ* index, in which SD is replaced by IQ, is suitable for any data distributions (*RPIQ*, $RPIQ=IQ/SEP$). The related chemometric indicators were calculated through the following relationships:

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{n}} \tag{11}$$

$$BIAS = \frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)}{n} \tag{12}$$

$$SEP = \sqrt{\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i-\mathrm{BIAS})^2}{n}} \tag{13}$$

3. Results and discussion

3.1. Quantitative analysis of baicalin by HPLC method

Fig. 1 shows a typical HPLC chromatogram of baicalin reference standard and a sample of Yinhuang oral solution. The retention time of the baicalin in the Yinhuang oral solution was the same with the reference standard. The calibration curve of the HPLC method was investigated before the real sample analysis. The calibration curve exhibited good linearity (R^2=0.9990) within the baicalin content range (0.051–0.450 μg). It can be concluded that the HPLC method satisfied the demand of quantitative analysis. Therefore, the reference values were accurate and could be used in NIR calibration models.

3.2. Chemometric model and model selection using usual chemometric indicators

3.2.1. PLS model

It is generally known that the spectral pre-treatments and the number of latent factors are critical to obtain a quantitative model. To predict the baicalin concentration profile, the calibration model referred to as PLS was built based on different spectral preprocessing treatments. The internal cross validation was used with a segment size of five. As seen from Table 2, raw spectra

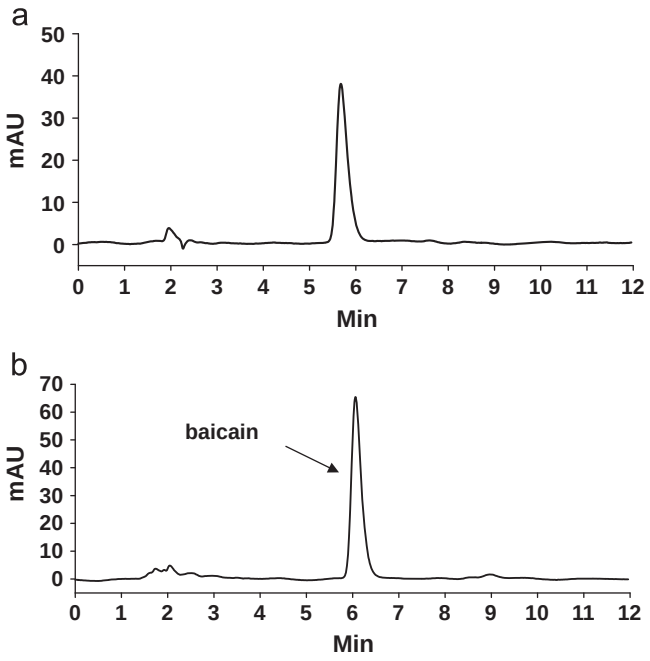

Fig. 1. Chromatograms of baicalin reference standard (a) and Yinhuang oral solution (b).

Table 2
Result of the PLS model with different spectral pre-processing methods.

Pretreatment	Latent factors	Calibration set		Validation set	
		R^2	RMSE	R^2	RMSE
Raw	1	0.9908	194.68	0.9889	217.2
1D	7	0.6473	1205	0.1425	1906.9
2D	3	0.4139	1553.6	0.1795	2244.1
SG	1	0.9908	194.4	0.9870	232.6
MSC	1	0.9907	195.8	0.9871	231.2
1D+SG	5	0.6455	1208.2	0.0814	1957.3

Raw: raw spectra, 1D: first derivative, 2D: second derivative, MSC: multiplicative signal correction, and SG: Savitzky–Golay.

were obviously superior to those of other spectral preprocessing methods for the PLS model. The calibration gave *RMSE* and R^2 values of 194.68 μg/mL and 0.9908, respectively. In the validation process, the *RMSE* and R^2 respectively were 217.2 μg/mL and 0.9889.

3.3. iPLS model

iPLS algorithm was applied in model calibration and the raw spectrum was divided into different sub-intervals. The optimal interval numbers were selected according to the lowest *RMSE*. As seen in Table 3, the number of interval selected was 38. It could be observed that interval number 11 provided models with best model performance compared with other models. The selection of spectral sub-intervals corresponded to 950–1005 nm.

3.3.1. BiPLS model

As for the BiPLS model, the data set was also split into different intervals. The most efficient interval number was 28, and models were calculated with each interval left out (the BiPLS result of other interval numbers is shown in the Supporting information). The first left out interval was the number 26 and this gave the poorest model with respect to *RMSE*. This procedure was continued until one interval remained (interval 1). The results are presented in Table 4. From Table 4, the *RMSE* of models decreased at first, and then increased as intervals were left out continually. It could be observed that the BiPLS model with best model performance was based on 4 intervals which were interval 11, 18, 13 and 1, with the *RMSE* as 92.5 μg/mL.

3.3.2. MWPLS model

MWPLS can provide informative regions and the approximate latent factors. The informative regions can be optimized by different moving window sizes. The moving window *H* varied from 13 to 41. The result demonstrated that the given window size has no significant effect on the selection of informative regions (data shown in Supporting information). However, a small window size shows weaker interference of the backward points and is expected to provide more accurate information about the position of the informative region than a big window size. Therefore, moving window size was selected as 13. From Fig. 2, the selection variables corresponded to 715–745 nm and 953–957 nm, and *RMSE* of the calibration model was 109.80 μg/mL.

3.3.3. Model selection using usual chemometric indicators

Table 5 describes prediction regressions for four multivariate models (the results of the correlation diagram between NIR predictions and HPLC results are shown in Supporting information). The optimum model was determined by the R^2, *RMSEP*, *RPD*, *RPIQ*, etc. The BiPLS model performance was the best among four models according to *RMSEP* and R^2 indicators; and the iPLS model was a little worse. However, if RPIQ parameter was selected as evaluation indicator, the MWPLS model performance was better than that of the BiPLS model. Sequentially, an interesting result was found that *RPIQ* value of the iPLS model was higher than that of the PLS model, despite the worse *RMSEP* and R^2 values. Therefore, above chemometric indicators reflected the integral performance of each model, but it did not fully exhibit accurate assessment on each concentration level of sample sets.

Table 3
Result of the iPLS Model with different sub-interval numbers.

Interval numbers	Selected interval	Latent factors	Calibration set		Validation set	
			R^2	*RMSE*	R^2	*RMSE*
20	6	4	0.9953	138.27	0.9952	142.69
22	6	3	0.9953	137.93	0.9948	146.17
24	7	4	0.9959	129.60	0.9954	137.40
26	7	3	0.9950	142.94	0.9949	146.92
28	8	2	0.9955	135.84	0.9955	137.37
30	8	3	0.9945	151.18	0.9944	156.82
32	9	3	0.9887	215.64	0.9869	233.15
34	10	3	0.9937	161.22	0.9931	168.31
36	10	4	0.9960	128.83	0.9957	133.40
38	11	4	0.9963	122.75	0.9961	127.18
40	11	4	0.9962	124.50	0.9960	130.40

Table 4
Selection of the most efficient interval regions through BiPLS.

Intervals numbers	Selected interval	RMSE	Number of variables
28	26	138.07	4200
27	28	131.41	4050
26	27	127.41	3900
25	25	122.98	3750
24	4	119.14	3600
23	21	118.52	3450
22	22	118.06	3300
21	5	116.60	3150
20	3	116.19	3000
19	24	115.86	2850
18	23	114.93	2700
17	20	110.84	2550
16	19	107.05	2400
15	14	106.78	2250
14	15	106.50	2100
13	2	105.95	1950
12	16	100.44	1800
11	6	94.44	1650
10	9	94.31	1500
9	7	94.25	1350
8	8	94.17	1200
7	10	94.09	1050
6	12	92.64	900
5	17	92.89	750
4	11	92.50	600
3	18	94.48	450
2	13	94.85	300
1	1	100.11	150

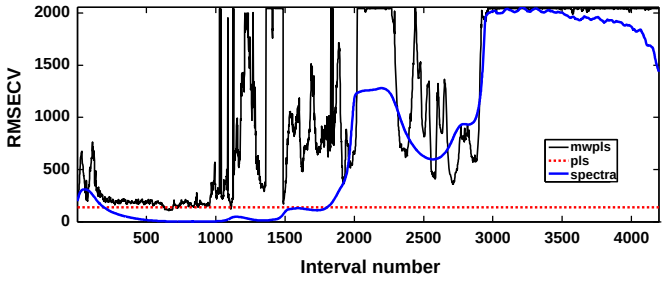

Fig. 2. Cross-validate prediction errors (*RMSECV*) for 13 window size and 7 latent variables for the full-spectrum PLS model.

3.4. Model selection based on total error concept

Accuracy profile (total error concept) was used to evaluate the model performance at each concentration level, as shown in Fig. 3. The acceptance limits were set at $\pm 10\%$ while the maximum risk to obtain results outside these acceptance limits was set at 5%. LLOQ in each model is obtained by calculating the smallest concentration beyond which the accuracy limits or β-expectation limits exceed the acceptance limits.

The PLS model was relatively unstable and points move toward the line of acceptance limits as concentration levels ranged from

Table 5
Statistics from different models.

Models	R^2	*RMSEP*	*Min*	*Max*	*Mea*	*Med.*	*Q1*	*Q3*	*SD*	*Skew*	*Kurtosis*	*RPIQ*
PLS	0.9918	152.85	1356	6544	4159	4733	2384	5748	1678	−0.32	−1.27	22.1
iPLS	0.9760	261.66	1283	6600	3968	4401	2061	5783	1791	−0.10	−1.34	27.1
BiPLS	0.9947	122.63	1403	6638	4136	3661	2162	5900	1772	−0.21	−1.36	33.2
MWPLS	0.9945	125.25	1297	6625	4118	4549	2241	5889	1730	−0.11	−1.34	34.8

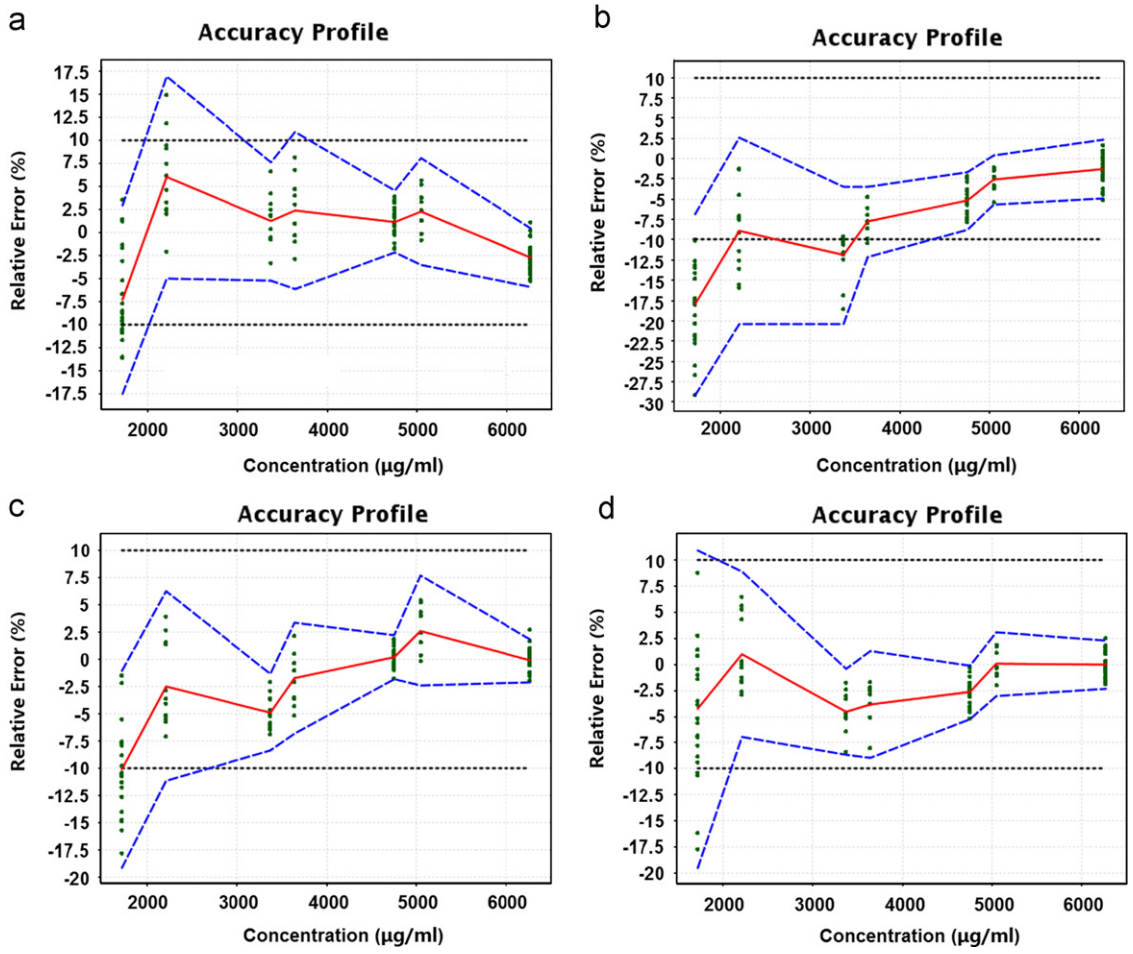

Fig. 3. Accuracy profiles of four models: (a) PLS, (b) iPLS, (c) BiPLS and (d) MWPLS. The plain line is the relative bias. The dashed lines are the β-expectations tolerance limits (β=95%) and the dotted lines represent the acceptance limits (± 10%).

1.0 mg/mL to 3.5 mg/mL. The bias becomes larger at intersection of 3.81 mg/mL. As for the iPLS model, it showed the largest bias compared with other models. The LLOQ was 4.34 mg/mL. And for BiPLS and MWPLS models it showed reduced relative error and indeed it can be observed that the β-expectations tolerance limits were included within the ± 5% acceptance limits at 3.37 mg/mL and 3.64 mg/mL concentration levels for two models. Furthermore, the MWPLS model had the lowest LLOQ, which was 2.08 mg/mL. Therefore, the MWPLS model was the most adequate one for Yinhuang oral solution.

In addition, Table 6 shows the ICH *Q*2(R1) validation criteria. The analytical properties, accuracy, precision and uncertainty could be observed for each model. The MWPLS model was still the most accurate one. The relative bias value was located at a maximum of −4.27% around the measured result. The risk value even attached to 0 at several concentration levels. Compared with the MWPLS model, the accuracy deteriorated in the iPLS model and the PLS model, especially at low concentration level. Certainly, the result indicated that each model provides good precision result. The overall difference of precision values was not obvious.

Finally, the linear profiles of the prediction models are shown in Fig. 4. For the PLS model, the regression equation was expressed as $Y=58.75+0.9803X$ with $R^2=0.9923$; for the iPLS model, the regression equation was $Y=-411.2+1.047X$ with $R^2=0.9960$; for the BiPLS model, the regression equation was $Y=-204.9+1.038X$ with $R^2=0.9973$; and for the MWPLS model, the regression equation was $Y=-109.5+1.011X$ with $R^2=0.9969$. It is well known that the slope and intercept closing to 1 and 0 respectively confirm the absence of proportional and constant total error of the model. The linearity results considering slope and intercept value in each multivariate calibration model demonstrated that the MWPLS model was still the best model for Yinhuang oral solution.

Therefore, a crucial point is to ensure that total error-based approaches should be considered in model selection strategy. Accuracy profile allows us to perfectly select reliable model by β-expectation tolerance interval.

Table 6
ICH Q2(R1) validation criteria in each calibration model.

Models	*LEV*	Accuracy			Precision		Uncertainty	
		REB	*RTL*	*RIS*	*REP*	*INP*	*UB*	*REU*
PLS	1714	−7.23	[−17.44, 2.98]	29.07	4.83	4.83	16.90	9.87
	2203	5.99	[−5.00, 16.98]	22.34	4.79	4.79	30.44	9.96
	3372	1.24	[−5.18, 7.66]	0.78	2.50	2.72	30.35	5.74
	3642	2.42	[−6.12, 10.96]	4.24	2.84	3.46	46.84	7.39
	4752	1.20	[−2.11, 4.52]	0.001	1.57	1.57	15.22	3.20
	5046	2.25	[−3.53, 8.04]	0.92	1.60	2.22	45.31	4.79
	6274	−2.68	[−5.89, 0.52]	0.002	1.52	1.55	18.63	3.15
iPLS	1714	−17.97	[−29.12, −6.81]	92.35	5.28	5.28	18.46	10.77
	2203	−8.92	[−20.45, 2.62]	42.21	5.02	5.02	31.95	10.46
	3372	−11.92	[−20.40, −3.45]	70.68	1.88	3.05	44.51	6.65
	3642	−7.78	[−12.11, −3.44]	14.14	1.89	1.89	19.86	3.93
	4752	−5.19	[−8.76, −1.62]	0.53	1.69	1.69	16.40	3.45
	5046	−2.62	[−5.70, 0.45]	0.02	1.22	1.31	21.43	2.76
	6274	−1.32	[−4.93, 2.30]	0.001	1.70	1.74	21.47	3.55
BiPLS	1714	−10.12	[−19.15, −1.08]	51.05	4.28	4.28	14.95	8.73
	2203	−2.52	[−11.21, 6.18]	4.68	3.79	3.79	24.09	7.88
	3372	−4.92	[−8.41, −1.43]	0.42	1.52	1.52	14.80	3.16
	3642	−1.78	[−6.88, 3.32]	0.27	2.01	2.17	25.84	4.57
	4752	0.17	[−1.82, 2.17]	0.00	0.88	0.93	10.97	1.92
	5046	2.59	[−2.47, 7.65]	0.54	1.58	2.01	39.05	4.31
	6274	−0.17	[−2.12, 1.79]	0.00	0.95	0.95	9.925	1.93
MWPLS	1714	−4.27	[−19.44, 10.91]	24.27	5.10	6.50	38.83	13.77
	2203	0.95	[−6.96, 8.85]	1.98	3.44	3.44	21.90	7.17
	3372	−4.57	[−8.67, −0.47]	0.71	1.79	1.79	17.37	3.72
	3642	−3.87	[−8.98, 1.25]	1.16	2.23	2.23	23.43	4.64
	4752	−2.67	[−5.22, −0.11]	0.00	1.16	1.20	13.34	2.46
	5046	0.01	[−3.00, 3.02]	0.00	1.27	1.30	19.79	2.72
	6274	−0.05	[−2.34, 2.25]	0.00	1.11	1.11	11.65	2.26

LEV: Level (μg/mL); REB: Relative bias (%); REP: Repeatability (RSD %); INP: Intermediate precision (RSD %); RTL: Relative expectation tolerance limits (%); *RIS*: risk (%); *UB*: uncertainty of the bias (μg/mL); *REU*: relative expanded uncertainty (%).

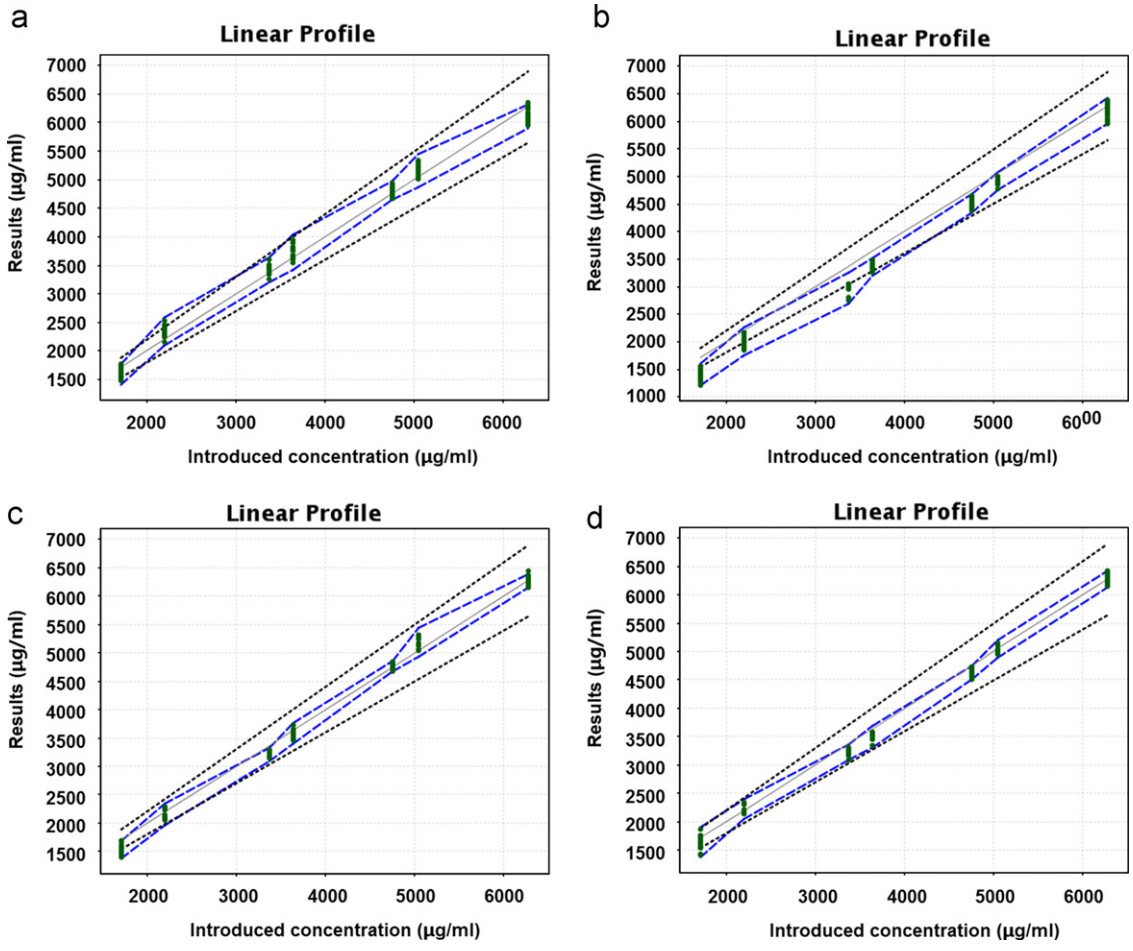

Fig. 4. Linear profile of different PLS models. The dashed limits in this graph correspond to the accuracy profile, i.e. the β-expectation tolerance limits expressed in absolute values. The dotted curves represent the acceptance limits at $\pm 10\%$ expressed in the concentration unit. The continuous line is the identity line $y=x$.

4. Conclusion

Throughout this paper, we proposed to conduct model selection base on the total error concept. Based on chemometric indicators and accuracy profile, different PLS models (PLS, iPLS, BiPLS and MWPLS) were assessed to quantify baicalin content in multi-component matrix. Usual chemometric indicators cannot fully assess the model ability in each low content range. Further, using the calculated tolerance interval, analytical methodology parameters i.e. linearity, relative bias, uncertainty, repeatability, intermediate precision, the LLOQ and risk were calculated to assess the model performance at each concentration level. In conclusion, a promising analysis of model selection strategy based on the total error concept (accuracy profile theory) was highly recommended to develop more robust model.

The strategy presented in this article is limited to lab data, and should be extended to manufacture data of CHM, including different low-concentration calibration sets. Additional research efforts are necessary to thoroughly evaluate the advantage of the total error-based approaches in model selection application, and to identify all its strengths and limitations.

Acknowledgments

This work was financially supported from the Ministry of Science and Technology of China Major Special Project "Significant Creation of New Drugs" (No: 2010ZX09502-002) and Innovation Team Foundation of Beijing University of Chinese Medicine, Beijing (2011-CXTD-11, Research Centre of TCM-information Engineering).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.talanta.2012.12.057.

References

[1] B. Zai, X. Shi, Y. Qiao, J. Chin. Mater. Med 35 (2010) 699–702.
[2] X.R. Zhu, N. Li, X.Y. Shi, Y.H. Qiao, Z.Y. Zhang, Chem. J. Chin. U 29 (2008) 906–911.
[3] W.L. Li, L.H. Xing, L.M. Fang, J. Wang, H.B. Qu, J. Pharm. Biomed. Anal 53 (2010) 350–358.
[4] W.L. Li, L.H. Xing, Y. Cai, H.B. Qu, Vib. Spectrosc. 55 (2011) 58–64.
[5] B. Xu, Z. Wu, Z. Lin, Chenglin Sui, X. Shi, Y. Qiao, Anal. Chim. Acta 720 (2012) 22–28.
[6] W.F. McClure, J. Near Infrared Spectrosc. 11 (2003) 487–518.
[7] S. Wold, M. Josefson, J. Gottfries, A. Linusson, J. Chemometr. 18 (2004) 156–165.
[8] L. Norgaard, A. Saudland, J. Wagner, J.P. Nielsen, L. Munck, S.B. Engelsen, Appl. Spectrosc. 54 (2000) 413–419.
[9] Y.X. Li, X.B. Zou, Y. Dong, Spectrosc. Spectral Anal 27 (2007) 2001–2004.
[10] X.B. Zou, J.W. Zhao, X.Y. Huang, Y.X. Li, Chemom. Intell. Lab. Syst 87 (2007) 43–51.
[11] A. Durand, O. Devos, C. Ruckebusch, J.P. Huvenne, Anal. Chim. Acta 595 (2007) 72–79.
[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] V. Bellon-Maurel, E. Fernandez-Ahumada, B. Palagos, J.M. Roger, A. McBratney, Trends Anal. Chem 29 (2010) 1073–1081.
[14] E. Rozet, A. Ceccato, C. Hubert, E. Ziemons, R. Oprean, S. Rudaz, B. Boulanger, P. Hubert, J. Chromatogr. A 1158 (2007) 111–125.
[15] P. Hubert, J.J. Nguyen-Huu, B. Boulanger, E. Chapuzet, P. Chiap, N. Cohen, P.A. Compagnon, W. Dewe, M. Feinberg, M. Lallier, M. Laurentie, N. Mercier, G. Muzard, C. Nivet, L. Valat, J. Pharm. Biomed. Anal. 36 (2004) 579–586.
[16] P. Hubert, J.J. Nguyen-Huu, B. Boulanger, E. Chapuzet, P. Chiap, N. Cohen, P.A. Compagnon, W. Dewe, M. Feinberg, M. Lallier, M. Laurentie, N. Mercier, G. Muzard, C. Nivet, L. Valat, E. Rozet, J. Pharm. Biomed. Anal. 45 (2007) 70–81.
[17] P. Hubert, J.J. Nguyen-Huu, B. Boulangerc, E. Chapuzet, N. Cohen, P.A. Compagnon, W. Dewe, M. Feinberg, M. Laurentie, N. Mercier, G. Muzard, L. Valat, E. Rozet, J. Pharm. Biomed. Anal. 45 (2007) 82–96.
[18] A.G. Gonzalez, M.A. Herrador, Trends Anal. Chem 26 (2007) 227–238.
[19] A. Bouabidi, E. Rozet, M. Fillet, E. Ziemons, E. Chapuzet, B. Mertens, R. Klinkenberg, A. Ceccato, M. Talbi, B. Streel, A. Bouklouze, B. Boulanger, P. Hubert, J. Chromatogr. A 1217 (2010) 3180–3192.
[20] M. Feinberg, J. Chromatogr. A 1158 (2007) 174–183.