Online Detection for Notoginseng Ingredients in Macroporous Resin Purification Process Using Near Infrared Technology in Combination with Partial Least Squares

R.X. Zhong[1], F. Long[1], D.S. Lin[2], H.R. Yang[2] and C.J. Wu[1,*]

[1]School of Pharmacy, Chengdu University of Traditional Chinese Medicine, Chengdu, P.R. China
[2]Chengdu Huasun Group Inc., Ltd., Chengdu, P.R. China

*Corresponding author: Tel/Fax: +86 02861801001; E-mail: wuchunjie@gmail.com

In this work, we have developed new prediction models for online detection of notoginseng ingredients (Ginsenoside Rg1, Ginsenoside Re and Notoginsenoside R1) during the macroporous resin purification process by near infrared technology in combination with partial least squares. In the experiments, calibration models were generated based on min max normalization (MMN) for Ginsenoside Rg1, Ginsenoside Re and First derivative + vector normalization for notoginsenoside R1. Finally, the root mean square errors of cross-validation (RMSECV) of Rg1 Re and R1 were 0.952, 0.106, 0.214 mg/mL, respectively; the root mean square errors of prediction (RMSEP) were 1.132, 0.121 and 0.381 mg/mL, respectively; the correlation coefficients (R^2) of cross-validation were 0.9858, 0.9787 and 0.9793, respectively; the correlation coefficients (R^2) of prediction were 0.9884, 0.9898 and 0.9709, respectively. In conclusion, this research indicated that NIR technology can be used for rapid analysis and online quality control of traditional Chinese medicine (TCM) in production process.

Keywords: Online detection, Near infrared technology, Notoginseng ingredients, Partial least squares.

INTRODUCTION

Traditional Chinese medicine (TCM) has attracted enormous attention among researchers all over the world, mainly attributable to its unique theory and long historical clinical practice[1,2]. As an important traditional Chinese medicine, Notoginseng belongs to Araliaceae is commonly used for promoting the blood circulation and removing the blood stasis. A significant extract from notoginseng is notoginseng triol saponins which, mainly composed of Ginsenoside Re, Ginsenoside Rg1 and Notoginsenoside R1, frequently contribute to cardiovascular and cerebrovascular thromboembolic disorders. Generally, the production process of notoginseng triol saponins mainly includes the percolation process and the purification process *via* macroporous resin[3]. To control the quality of the purification process online is a critical point in the entire production process. Currently, high performance liquid chromatography (HPLC) method is widely used to control the quality of macroporous resin purification process by determining the contents of Ginsenoside Rg1, Ginsenoside Re and Notoginsenoside R1. However, this method suffer from time-consuming and destructiveness[4,5]. This appears to be a strong motivation to develop a new method which can achieve fast acquisition, non-destructive characteristics for online quality control in the macroporous resin purification process.

In recent years, near infrared technology, due to its rapid analysis, the characteristics of the non-invasion and non-destructiveness and the accuracy for multi-parameters, has led to a significant interest and proved to be a powerful tool for qualitative and quantitative analysis in a variety of applications such as food[6], agricultural[7,8] and pharmaceutical industries field[9-11]. The most prominent absorption bands of near infrared consist of overtones and combinations of fundamental vibrations displayed by -CH, -NH, -OH groups[12,13] and the principles of near infrared spectroscope (near infrareds) are intrinsically different from the conventional analytical techniques such as HPLC[14,15]. To obtain the meaningful correlations from near infrareds, an appropriate chemometrics method is usually required. As an effective statistical method, partial least squares has become the most popular chemometrics method since it requires fewer latent variables and can handle nonlinearities better than other methods[16]. Despite the considerable potential applications of near infrareds, few studies, however, have been reported on determining the ingredients of notoginseng in macroporous resin purification process by exploiting near infrareds in combination with partial least squares.

In this paper, new prediction models of Ginsenoside Rg1, Ginsenoside Re and Notoginsenoside R1 were proposed for online analysis during the macroporous resin purification

process by combing with near infrareds and partial least squares. Specifically, these models were optimized by cross-validation; meanwhile, the root mean square error of prediction (RMSEP) and the correlation coefficient (R2) in prediction set were chosen to use for evaluating them. On the above foundation, the macroporous resin purification process can be monitored online by determining the contents of Ginsenoside Rg1, Ginsenoside Re and Notoginsenoside R1.

EXPERIMENTAL

All samples were supplied by Sichuan Huasheng pharmaceutical Co. Ltd. The production process was divided into four steps as follows. First, ten batches of Notoginseng were crushed into powders by a cyclone mill. Second, these powders were moistened with 60 % ethanol and put in a closed vessel over 24 h, after that, we percolated these powders and collected the leachate. Third, the leachate was poured into the macroporous resin and was eluted with water for 2.5 h. Fourth, the leachate was eluted again with 40 % ethanol for 6 h and each of 25 mL leachate as a sample was collected in the beaker every 15 min. Eventually, a total of 240 liquid samples, in which each batch were composed of 24 samples, could be directly detected by PbS detector at a room temperature.

Ginsenoside Rg1 (batch number: 110703-200626), Notoginsenoside R1 (batch number: 110745-200415), Ginsenoside Re (batch number: 110754-200421) standards were purchased from the National Institute for Institutes for Food and Drug Control (Beijing, China). HPLC-grade acetonitrile was obtained from the Tianjin Kermel Chemical Reagent Company (Tianjing, China). Water was purified by an ultrapure water instrument. All other reagents were analytical grade.

Detection method: Near infrared spectra were measured by a Bruker Matrix-I Fourier-transform near infrared spectrometer (FT-NIR) equipped with a PbS detector, sample cup and rotary tables. The system was operated by OPUS spectral acquisition and processing software (Bruker Optik, Ettlingen, Germany). The result of the scans was the mean of 64 scans which were measured with 8 cm^{-1} resolution over the wavenumber range from 12500 to 3600 cm^{-1} and the air absorbance was regarded as the reference standard. Based on the above foundation, we collected the samples in beakers and then inserted the probe into these samples. Every sample was scanned three times and the average spectrum of each sample was used for the final analysis.

The pretreatment can eliminate the influence of confounding factors, filter spectral noise, optimize spectral region and get more effective information. The OPUS software contained 11 important pretreatment methods. In this paper, partial least squares regression was used to develop calibrations and validate internally by cross-validation.

An Agilent 1200 high performance liquid chromatography (HPLC) system, consisting of quaternary pump, autosampler, UV-visible detector at a wavelength of 210 nm and thermostat maintained at 30 °C, was used to separate and analyze 10 μL sample injections over a waters symmetry shield RP18 (250 × 3.9 mm, 5 μm). The elution system was composed of acetonitrile-water (19.5:80.5) and separations were carried out at the flow of 1 mL/min. The 250 mg Ginsenoside Rg1, 40 mg Ginsenoside Re and 80 mg Notoginsenoside R1 were dissolved in 100 mL mobile phase by ultrasonic processing at room temperature. Then the standards and the samples of macroporous resin purification were filtered by microporous membrane with a pore size of 0.45 μm at room temperature.

RESULTS AND DISCUSSION

Total 240 samples of ten batches were collected in the beakers. Among the ten batches, we chose eight batches of 192 samples as calibration set, whereas the remaining two batches of 48 samples were set as validation set. During the macroporous resin purification process, the contents of Notoginsenoside R1, Ginsenoside Rg1 and Ginsenoside Re were varied with time and the results were shown in Table-1.

Primary selection of the spectrum region: The spectrum of samples was shown in Fig. 1. As can be seen from this figure, there was no distinct absorbing in the spectrum over the wavelength range of 12000-8000 cm^{-1}. Therefore, measurements were carried out in a range of 8000-4000 cm^{-1} due to the fact that this region had abundant chemical information and composition content of samples. Furthermore, this selection of the spectrum region can reduce the computational cost and prevent the overfitting problem.

Selecting the pretreatment methods: The OPUS software (Bruker Optik, Ettlingen, Germany) developed by Bruker Instruments LTD contains 11 important spectral pretreatment methods: Untreated, Straight Line Subtraction (SLS), Constant Offset Elimination (COE), Min Max Normalization (MMN), multiplicative scatter calibration (MSC), Vector normalization (VN), First Derivative, Second Derivative, First Derivative +

TABLE1
CONTENTS OF 10 BATCHES SAMPELS

No.	Notoginsenoside R1 (mg/mL)	$\overline{x} \pm s$	Ginsenoside Rg1 (mg/mL)	$\overline{x} \pm s$	Ginsenoside Re (mg/mL)	$\overline{x} \pm s$
1	0.132-4.753	1.024 ± 1.398	0.121-9.064	4.441 ± 3.848	0.012-1.753	0.491 ± 0.526
2	0.091-4.412	1.536 ± 1.459	0.942-8.526	4.803 ± 3.728	0.075-1.803	0.623 ± 0.530
3	0.123-4.832	1.653 ± 1.536	0.791-16.721	5.802 ± 4.294	0.137-2.067	0.717 ± 0.546
4	0.111-4.040	1.041 ± 1.146	0.421-16.038	4.895 ± 4.700	0.123-1.745	0.539 ± 0.493
5	0.105-4.037	1.780 ± 1.431	0.465-17.200	6.033 ± 5.058	0.106-1.763	0.617 ± 0.566
6	0.103-5.714	2.273 ± 2.257	0.959-16.733	7.146 ± 6.018	0.074-2.148	0.799 ± 0.776
7	0.108-4.836	1.427 ± 1.374	0.894-14.538	4.704 ± 3.783	0.100-1.752	0.559 ± 0.474
8	0.139-4.174	1.126 ± 1.278	0.279-10.673	4.871 ± 3.459	0.014-1.366	0.532 ± 0.427
9	0.112-4.943	1.275 ± 1.355	0.171-14.814	4.285 ± 3.817	0.073-2.133	0.555 ± 0.549
10	0.106-4.399	1.305 ± 1.281	0.738-14.583	5.025 ± 3.766	0.106-1.802	0.603 ± 0.479
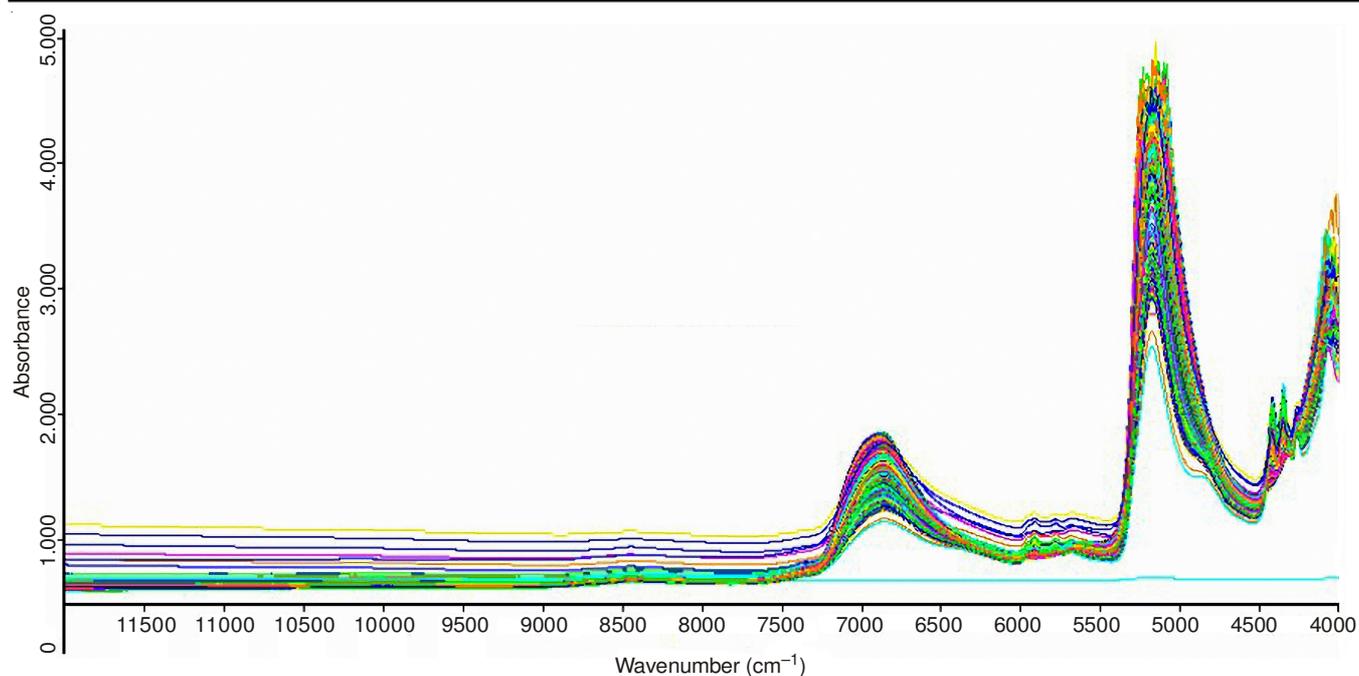
Fig. 1. Spectrum of the samples

TABLE-2
COMPARISON OF VARIOUS PRETREATMENT METHODS

	Pretreatments	D	Calibration set		Validation set	
			R^2	RMSECV	R^2	RMSEP
Ginsenoside Rg1	MMN	14	0.9817	0.958	0.9884	1.321
	MMN	15	0.9858	0.952	0.9884	1.132
	MSC	14	0.9785	0.933	0.9834	1.524
	MSC	15	0.9809	0.966	0.9811	1.523
Ginsenoside Re	MMN	10	0.9787	0.106	0.9898	0.121
	MMN	11	0.9792	0.104	0.9804	0.173
	MSC	10	0.9781	0.107	0.9805	0.172
	MSC	11	0.9758	0.113	0.9774	0.183
Notoginsenoside R1	MSC	10	0.9667	0.271	0.9685	0.467
	MSC	11	0.9618	0.288	0.9643	0.515
	First Derivative + VN	10	0.9793	0.214	0.9709	0.381
	First Derivative + VN	11	0.9593	0.297	0.9716	0.419

Straight Line Subtraction, First Derivative + Vector normalization and First derivative + multiplicative scatter calibration. The results of the comparison of various pretreatment methods for the three ingredients were shown in Table-2.

Through the OPUS, the assessment of the predictive ability and the robustness of different models depended on these statistical parameters such as low root mean square error of calibration (RMSEC), prediction (RMSEP) and high correlation coefficient (R^2). As shown in Table-2, MMN was most fit for Ginsenoside Rg1 and Ginsenoside Re, whereas First Derivative + VN was most fit for Notoginsenoside R1.

Determining the dimension (D) of partial least squares factors: It was well known that the dimension of partial least squares factors was a critical parameter for partial least squares with its application to develop a calibration model. Even in the same spectral pretreatment method, different value of dimension of partial least squares factors would directly influence the values of R^2, RMSECV and RMSEP. For example, some useful and sufficient information may not be extracted from the spectrum when dimension of partial least squares

factors was too small, whereas the actual information might be over-fitting or misrepresented with a larger value of dimension of partial least squares factors. Fortunately, the OPUS can successfully recommend the ideal value of dimension of partial least squares factors. The statistical parameters *vs.* the dimension were depicted in Fig. 2.

It was observed from Fig. 2 that with the increment of dimension of partial least squares factors the value of RMSECV decreased but R^2 increased; however, the two parameters tended to be stable when dimension of partial least squares factors achieved a certain value. Finally, OPUS recommended the dimensions 15, 10, 10, respectively, for Ginsenoside Rg1, Ginsenoside Re, Notoginsenoside R1.

Evaluating the predicted result of the validation set: Based on the optimal parameters obtained from previous subsections, we established the models of Ginsenoside Rg1, Ginsenoside Re, Notoginsenoside R1 *via* partial least squares algorithm for the calibration set. Then we applied the three models to predict the contents of the three ingredients of Notoginseng for validation set. The relationship between the
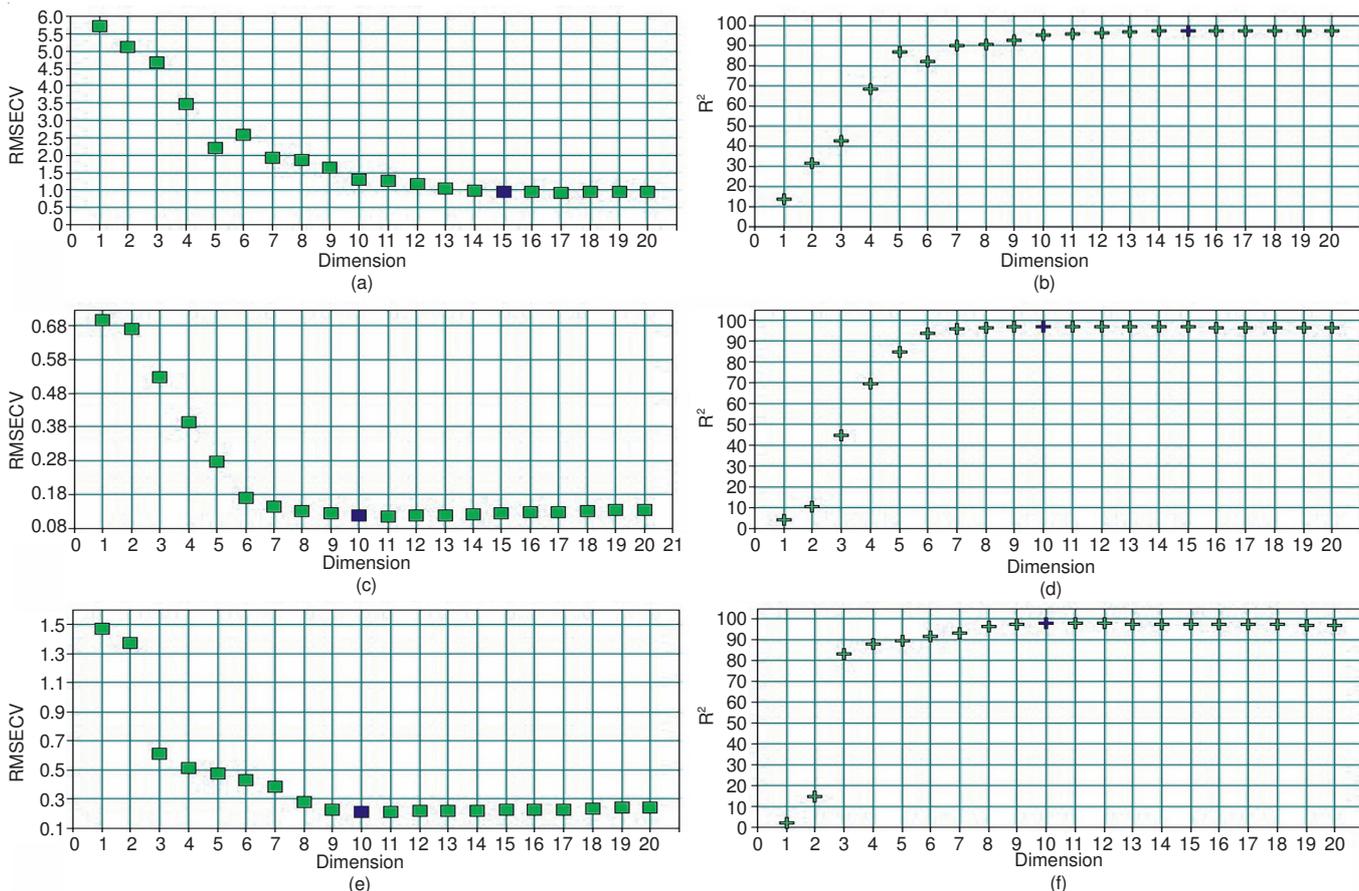
Fig. 2. RMSECV versus dimension (D) for (a) Ginsenoside Rg1, (c) Ginsenoside Re and (e) Notoginsenoside R1, respectively; R² *vs.* dimension (D) for (b) Ginsenoside Rg1, (d) Ginsenoside Re and (f) Notoginsenoside R1, respectively
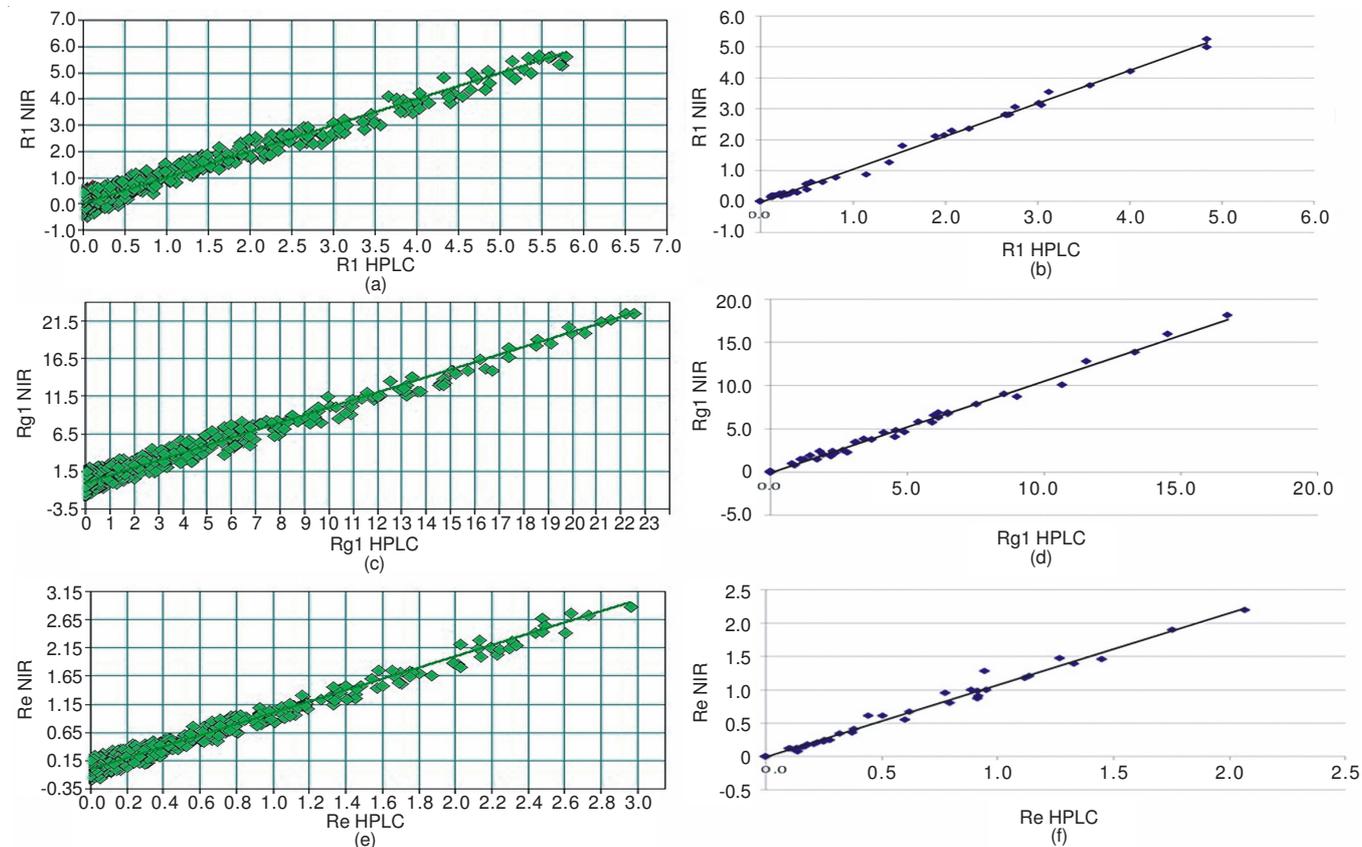
Fig. 3. Correlation between values by NIRs and HPLC for (a) Notoginsenoside R1, (c) Ginsenoside Rg1and (e) Ginsenoside Re of the calibration set; Correlation between values by NIRs and HPLC of (b) Notoginsenoside R1, (d) Ginsenoside Rg1and (f) Ginsenoside Re of the validation set

TABLE-3
PARAMETERS OF OPTIMAL MODELS FOR THREE ACTIVE INGREDIENTS

		Ginsenoside Rg1	Ginsenoside Re	Notoginsenoside R1
Pretreatment		MMN	MMN	First Derivative + VN
Wavelength (cm⁻¹)		6101.8-5449.9 4424.4-4246.5	6101.8-5449.9	6101.8-5449.9 4601.4-4246.5
Samples	Calibration	192	192	192
	Validation	48	48	48
D		15	10	10
RMSECV (mg/mL)		0.952	0.106	0.214
R^2 (the calibration)		0.9858	0.9787	0.9793
RMSEP (mg/ml)		1.132	0.121	0.381
R^2 (the validaition)		0.9884	0.9898	0.9709

corresponding content values measured by HPLC and near infrareds was showed in Fig. 3.

As can be seen from Fig. 3, there was no great difference between the predictive values measured by near infrareds and the values measured by HPLC. Furthermore, the content detection by near infrareds just required 20 seconds, while at least 30 min were required by HPLC. The parameters of the three optimal models were shown in Table-3.

Conclusion

This study has shown that near infrareds can be successfully applied to the online analysis for Notoginseng ingredients in macroporous resin purification process. By further combining partial least squares, the macroporous resin purification process can be controlled online to produce the best products. Overall, the proposed method based on near infrareds and partial least squares demonstrates a great promise for rapid analysis and online quality control in pharmaceutical industry.

REFERENCES

1. Y. Jiang, B. David, P.F. Tu and Y. Barbin, *Anal. Chim. Acta*, **657**, 9 (2010).
2. J. Xue, C. Wu, L. Wang, S. Jiang, G. Huang, J. Zhang, S. Wen and L. Ye, *Food Chem.*, **126**, 725 (2011).
3. Q.F. Xu, X.L. Fang and D.F. Chen, *J. Ethnopharmacol.*, **84**, 187 (2003).
4. C.C. Bai, S.Y. Han, X.Y. Chai, Y. Jiang, P. Li and P.F. Tu, *J. Liq. Chromatogr. Rel. Technol.*, **32**, 242 (2008).
5. W. Luo, J. Wu, X.K. Wang, X. Lin and H. Li, *Anal. Methods*, **5**, 1337 (2013).
6. J.A. Cayuela, *Postharvest Biol. Technol.*, **47**, 75 (2008).
7. J.W. Zhao, Q.S. Chen, X.Y. Huang and C.H. Fang, *J. Pharm. Biomed. Anal.*, **41**, 1198 (2006).
8. Q.S. Chen, J.W. Zhao, S. Chaitep and Z.M. Guo, *Food Chem.*, **113**, 1272 (2009).
9. M. Blanco, R.G. Bañó and E. Bertran, *Talanta*, **56**, 203 (2002).
10. Y. Chen, M.Y. Xie, Y. Yan, S.B. Zhu, S.P. Nie, C. Li, Y.X. Wang and X.F. Gong, *Anal. Chim. Acta*, **618**, 121 (2008).
11. D. Wu, J.Y. Chen, B.Y. Lu, L.N. Xiong, Y. He and Y. Zhang, *Food Chem.*, **135**, 2147 (2012).
12. G. Reich, *Adv. Drug Deliv. Rev.*, **57**, 1109 (2005).
13. J. Luypaert, D.L. Massart and Y. Vander Heyden, *Talanta*, **72**, 865 (2007).
14. W.L. Li, L.H. Xing, Y. Cai and H.B. Qu, *Vib. Spectrosc.*, **55**, 58 (2011).
15. H. Winning, N. Viereck, T. Salomonsen, J. Larsen and S.B. Engelsen, *Carbohydr. Res.*, **344**, 1833 (2009).
16. Y. Roggo, P. Chalus, L. Maurer, C. Lema-Martinez, A. Edmond and N. Jent, *J. Pharm. Biomed. Anal.*, **44**, 683 (2007).