**Natural Product Sciences**
24(3) : 164-170 (2018)
https://doi.org/10.20307/nps.2018.24.3.164

# Classficiation of Bupleuri Radix according to Geographical Origins using Near Infrared Spectroscopy (NIRS) Combined with Supervised Pattern Recognition

**Dong Young Lee[1], Kyo Bin Kang[2], Jina Kim[1], Hyo Jin Kim[3], and Sang Hyun Sung[1,*]**

*[1]College of Pharmacy and Research Institute of Pharmaceutical Science, Seoul National University, Seoul 151-742, Republic of Korea*
*[2]College of Pharmacy, Sookmyung Women's University, Seoul 04310, Republic of Korea*
*[3]College of Pharmacy, Dongduk Women's University, Seoul 136-714, Republic of Korea*

**Abstract** – Rapid geographical classification of Bupleuri Radix is important in quality control. In this study, near infrared spectroscopy (NIRS) combined with supervised pattern recognition was attempted to classify Bupleuri Radix according to geographical origins. Three supervised pattern recognitions methods, partial least square discriminant analysis (PLS-DA), quadratic discriminant analysis (QDA) and radial basis function support vector machine (RBF-SVM), were performed to establish the classification models. The QDA and RBF-SVM models were performed based on principal component analysis (PCA). The number of principal components (PCs) was optimized by cross-validation in the model. The results showed that the performance of the QDA model is the optimum among the three models. The optimized QDA model was obtained when 7 PCs were used; the classification rates of the QDA model in the training and test sets are 97.8% and 95.2% respectively. The overall results showed that NIRS combined with supervised pattern recognition could be applied to classify Bupleuri Radix according to geographical origin.

**Keywords** – Near infrared spectroscopy, Bupleuri Radix, Geographical classification, Supervised pattern recognition

## Introduction

Bupleuri Radix is the dried roots of *Bupleurum falcatum*, which belongs to the Umbelliferae, and is a commonly used medicinal herb for the treatment of fever, pain and inflammation associated with influenza and the common cold.[1] Recently, studies have shown that Bupleuri Radix has anti-inflammatory,[2] immunomodulatory,[3] anti-gastric ulcer,[4] and anti-influenza[5] activities.

*Bupleurum falcatum* is widely distributed in East Asia.[6] [6]However, the chemical composition of medicinal herbs is somewhat variable according to their growing conditions such as cultivation soil and climate based on geographical origins, even the herbs come from the same species. Therefore, a fast, precise and accurate analytical method to determine the geographical origins is required for the quality control of medicinal herbs.

Quality control of Bupleuri Radix has been studied using several analytical methods, including gas chromatography (GC),[7] high-performance liquid chromatography (HPLC),[8-12] high-performance thin layer chromatography (HPTLC), nuclear magnetic resonance (NMR),[13] and capillary zone electrophoresis (CZE).[14] These analytical methods are precise and accurate, but they are destructive, time-consuming, labor-intensive and costly. Furthermore, it is not enough to select only a few markers to reflect the therapeutic effects of numerous ingredients in medicinal herbs.[15]

Near-infrared spectroscopy (NIRS) has been shown to be a powerful analytical method for qualitative and quantitative analyses in the food, agricultural, and pharmaceutical industries. NIRS provides rapid, precise and non-destructive methods requiring minimal or no sample preparation[16] and the medicinal herbs do not lose their natural character due to NIRS.[17] In recent years, many studies have classified medicinal herbs according to their geographical origins by NIRS, such as Paeoniae Radix,[18] *Ganoderma lucidum*,[17] *Glycyrrhiza uralensis*,[19] *Camellia sinensis*,[20] Pseudostallariae Radix,[21] and Scrophulariae Radix.[22] However, geographical classification of Bupleuri Radix using NIRS has not been done yet.

Supervised pattern recognition refers to methods with

*Author for correspondence
Sang Hyun Sung, College of Pharmacy and Research Institute of Pharmaceutical Science, Seoul National University, Seoul 151-742, Republic of Korea.
Tel: +82-2-880-7859; E-mail: shsung@snu.ac.kr

the information about the category membership of the samples to a certain group (training set) in order to classify new unknown samples in one of the known categories (test set) based on its pattern of measurements.[23] The classification model is established by a training set of samples with categories. The model performance is assessed by means of samples in a test set. Several types of supervised pattern recognition methods are applied in chemistry, biology, pharmaceutical, and food science; however, they fundamentally differ in the way they obtain the classification results. Accordingly, selecting the most appropriate classification model is required in supervised pattern recognition methods.

In this study, three supervised pattern recognition methods were used to establish a robust classification model. These methods were the partial least square - discriminant analysis (PLS-DA), quadratic discriminant analysis (QDA) and radial function basis - support vector machine (RBF-SVM), respectively. Principal component analysis (PCA) was performed on the NIR spectra to remove the outliers and achieve some principal components (PCs) as the inputs of the QDA and RBF-SVM models. In addition, some parameters of the supervised pattern recognition methods were optimized by cross validation. Moreover, the model performance was assessed by means of classification rates of the training and test sets.

## Experimental

**Samples and reagents** – In the experiments, seventy-five Bupleuri Radix samples from two geographical origins in South Korea were collected by Dr. Wan Kyun Hwang, a professor at Chung-Ang University, and fifty-six samples from three geographical origins in China were collected by Dr. Young-Bae Seo, a professor at Daejeon University. Voucher specimens were deposited in the College of Pharmacy, Seoul National University. The samples were randomly divided into two subsets for the classification. Two-thirds of the samples called the training set were used to build the classification model, and the other third called the test set was used to assess the robustness of the classification model. Detailed information about the samples is listed in Table 1. After all the samples were cleaned and air-dried, they were milled into powder with a grinder for 30 seconds. To reduce the effect of particle size, the pulverized samples were sieved with a 100-mesh sieve (150 μm). This sieved powder was put into a glass vial with a stopper and dried for 12 hours in an oven at 50°C to remove the moisture in the samples before the NIR spectroscopy analysis.

**Table 1.** Summary of testing Bupleuri Radix samples

| Geographical origins | Number of samples | |
|---|---|---|
| | Training set | Test set |
| South Korea | | |
| Goheung | 35 | 17 |
| Whasun | 16 | 7 |
| China | | |
| Shanxi | 10 | 5 |
| Gansu | 17 | 8 |
| Shaanxi | 11 | 5 |

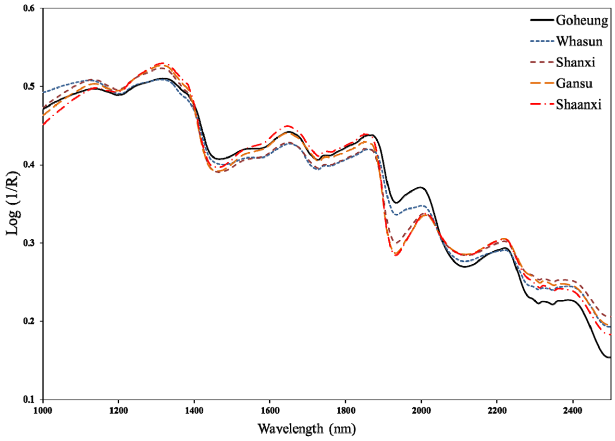


**Fig. 1.** Average NIR reflectance specrtra of Bupleuri Radix obtained from raw data.

**NIR spectroscopy measurements** – The near infrared reflectance spectra of the Bupleuri Radix powder were obtained with an NIR system (MPA; Bruker optics, Ettlingen, Germany) over a wavelength range of 1000 – 2500 nm, using 32 scans and 1.25 nm resolution per spectrum. Each spectrum was an average of 32 scanning spectra. About 0.8 g of each sample powder were individually transferred to a glass sample vial. The spectra were acquired in the reflectance mode with an empty glass sample vial as a reference standard. Each sample spectrum was measured three times and the final spectra were averaged.

**Spectra preprocessing** – Fig. 1, shows the raw mean spectra of the Bupleuri Radix samples from five geographical origins. Near infrared reflectance spectra are affected by both the chemical and physical properties of the samples, and the latter properties influence the majority of unwanted variance among the spectra. Therefore, it is necessary to perform spectral pre-processing to decrease the systematic noise, such as light scattering, path length differences, baseline variation and so on.[24] In this study, several spectral preprocessing methods were

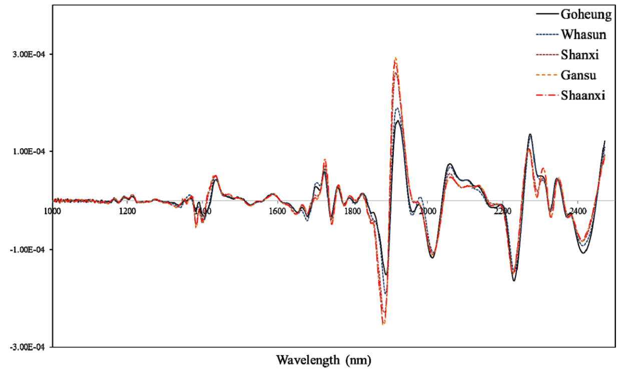


**Fig. 2.** Average NIR reflectance specrtra of Bupleuri Radix with 2nd derivative preprocessing.

used comparatively to obtain the optimum results. They include first derivative, second derivative, standard normal variate (SNV), and multiplicative scatter correction (MSC). To avoid the noise enhancement, which was the consequence of the derivative, the Savitzky-Golay smoothing filter (25 points in each sub-model and a distance of 2 between each point) was used. Compared with the results obtained by these preprocessing methods, the second derivative method is improved compared to the others. Therefore, the second derivative method was used in this work. The NIR spectra after second derivative preprocessing are shown in Fig. 2.

**Partial least squares discriminant analysis (PLS-DA)** – PLS-DA modeling is adimensionality reduction method for maximizing a relationship between dependent variables (Y) and independent variables (X).[25] The principle of PLS is to obtain the components in the X that explain as much as possible the relevant variations in the X and have a maximal correlation with the target value in Y, giving less weight to the variations that are unimportant or noisy. In PLS-DA, a ‘dummy’ Y consists of ‘0’ and ‘1’. The closer an element of a certain class is to ‘1’ and the elements of the other class of ‘0’, the more likely the element is a member of a particular class. A cut-off value is also needed so that a sample is assigned to a certain class. In this study, ±0.35 was used as the cut-off value.

**Quadratic discriminant analysis (QDA)** – QDA is closely related to linear discriminant analysis (LDA). LDA is based on the determination of linear discriminant functions, which maximize the ratio of inter-class variance and minimize the ratio of intra-class variance. Additionally, in LDA, it is assumed that the measurements from each class are supposed to follow a normal distribution and be linearly separated. LDA requires that the number of samples is higher than that of the variables; therefore, a variable reduction method such as PCA might be needed. Unlike LDA, QDA which develops parabolic boundaries and not linear boundaries, is less subjected to constraints in the distribution of objects in space compared to LDA.[23]

**Radial basis function–support vector machine (RBF-SVM)** – The SVM algorithm creates a hyperplane or set of hyperplanes in a high dimensional space, which can be used for classification. A good classification is achieved by the hyperplane that has the largest distance to the nearest data points of any class (functional margin), because generally, as the margin increases, the generalization error of the classifier decreases.[26] In SVM, a transformation into a high dimension feature space is implemented by a kernel function. Generally, there are three classical kernel functions: polynominal, sigmoid, and radial basis function (RBF) kernel function. Selection of the kernel function is very important for the performance of the SVM model. The RBF kernel function is the most common choice without prior experienced knowledge.[27] The RBF kernel function is shown in following equation.[27]

$$K(x_i,x_j)=\exp\left(\frac{-\|x_i-x_j\|^2}{\gamma^2}\right)$$

Here, $\gamma$ is the kernel parameter, which is the bandwidth of the RBF kernel function.

To achieve good performance, some parameters in the RBF-SVM model have to be optimized. These parameters include the following : (1) penalty parameter ($C$), which determines the trade-off between minimizing the model complexity and the training error, and (2) kernel parameter ($\gamma$), which is the bandwidth of the RBF kernel function. In this study, a “grid-search” on these parameters was used by cross-validation. It was found that the optimal RBF-SVM model is established when $C = 6$ and $\gamma = 129$.

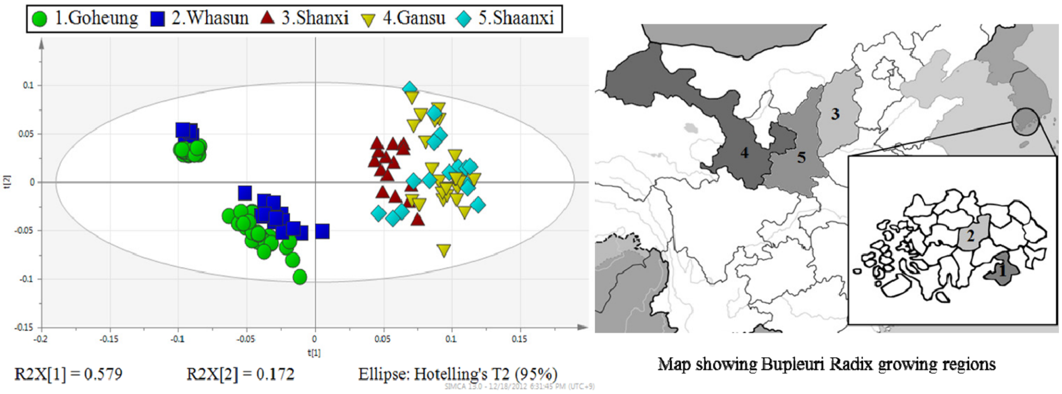
**Data analysis** – NIRS spectral data acquisition and spectral preprocessing were performed by the OPUS 7.0software (Bruker Optics, Germany). The SIMCA 13 software (Umetrics, Sweden) was used for PCA and PLS-DA. The data sets were in the Pareto scaling mode before PCA and PLS-DA. The Unscrambler X software (Camo, Norway) was used for the QDA and RBF-SVM.

## Result and discussion

**Spectra investigation** – As shown in Fig. 1, the NIR spectra of the Bupleuri Radix samples from South Korea and China have considerable differences. There are water absorption bands around 1933 nm according to two vibrations of O-H deformation and stretching. The most intensive band in the spectra belonged to the first overtone of N-H stretching of the amide bond (1468 nm), followed by the second overtone of C-H stretching ($CH_3$, 1195 nm), the first overtone of C-H stretching ($CH_2$, 1727 nm), the symmetric stretching of N-H (amide, 2119 nm), the stretching and deformation vibration of $CH_2$ and allylic $CH_2$ (2311 nm, 2348 nm) and the stretching of C-H and C-C (starch, 2497 nm).

**Principal component analysis (PCA)** – PCA is an unsupervised method of visualization and data compression widely used in NIRS technology. To visualize the cluster trends of the samples, a score plot was obtained using the main two principal components (Fig. 3). The main two principal components, PC1 and PC2, accounted for 75.1% of the data variance. As shown in Fig. 3, separation between the South Korean (Goheung, Whasun) and Chinese (Shanxi, Gansu, Shaanxi) samples was successfully performed by PC1,but a partial overlap was observed between the Goheung and Whasun samples. Especially, more overlap was observed among the Shanxi, Gansu and Shaanxi samples. However, only visual classification results were obtained by the PCA. For actual classification, four different classification methods, PLS-DA, QDA, and RBF-SVM were performed in the following studies.

**Classification by partial least square-discriminant analysis (PLS-DA) model** – When the PLS-DA model is established, one of the key issues is the selection of the ideal number of latent variables (LVs), which is generally performed based on a cross-validation.[28] In this study, selection of 5 LVs in terms of the classification performance was found to be the most suitable number of LVs. Table 2, shows the results of partial least square-discriminant analysis (PLS-DA) of Bupleuri Radix. The classification rate was 88.8%in the training set and 83.3% in the test set, respectively. It can be seen that the model is not accurate enough for geographical classification. In the cases of the samples from Goheung, Whasun and Gansu, a classification rate of 100% was obtained. However, the



Map showing Bupleuri Radix growing regions

**Fig. 3.** Two dimensional score plot of the top two principal components (PCs) for all samples.

**Table 2.** Classification results by PLS-DA model

| PLS-DA | | Goheung | Whasun | Shanxi | Gansu | Shaanxi | Classification rate (%) | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | per group | all groups |
| Training set (n=89) | Goheung | 35 | 0 | 0 | 0 | 0 | 100 | 88.8 |
| | Whasun | 0 | 16 | 0 | 0 | 0 | 100 | |
| | Shanxi | 0 | 0 | 8 | 0 | 2 | 80 | |
| | Gansu | 0 | 0 | 0 | 17 | 0 | 100 | |
| | Shaanxi | 1 | 0 | 0 | 7 | 3 | 27.3 | |
| Test set (n=42) | Goheung | 17 | 0 | 0 | 0 | 0 | 100 | 83.3 |
| | Whasun | 0 | 7 | 0 | 0 | 0 | 100 | |
| | Shanxi | 0 | 0 | 2 | 0 | 3 | 40 | |
| | Gansu | 0 | 0 | 0 | 8 | 0 | 100 | |
| | Shaanxi | 0 | 0 | 0 | 4 | 1 | 20 | |

* The rows indicate the true sample class and that the columns refer to the observed class.

classification rate of the samples from Shanxi and Shaanxi was 40% and 20% in the test set respectively.

As shown in Table 2, the PLS-DA model enabled a perfect classification between samples from Goheung and Whasun that were not possible with PCA. Nevertheless, the result indicated that part of the samples collected from Shanxi were classified as Shaanxi, and most of the samples collected from Shaanxi were classified as Gansu. Therefore, the classification between the samples from China was difficult using the PLS-DA.

**Classification by quadratic discriminant analysis (QDA) model** – The number of principal components (PCs) is important in the performance of the QDA classification model. The classification rates of the test set were used to optimize the number of PCs. The Fig. 4., shows the classification rate of the QDA model after a cross validation. The optimal QDA model was achieved when the number of PCs was 7. Table 3 shows the results of the quadratic discriminant analysis model of Bupleuri Radix. The classification rate was 97.8% in the training set and 95.2% in the test set, respectively. In particular, in the case of samples from Shanxi and Shaanxi, the QDA model showed a significant superiority over the PLS-DA models.

**Classification by radial basis function - support vector machine (RBF-SVM) model** – Because the above two models did not give a complete answer to the classification problem, the RBF-SVM model was used in this work. As previously indicated, two parameters, the penalty parameter ($C$) and kernel parameter ($\gamma$), were optimized according to the highest classification rate by a cross-validation. When $C = 6$ and $\gamma = 129$, the optimal RBF-SVM model was achieved. After the two parameters of the RBF-SVM model were optimized, the number of PCs was also determined. Fig. 4, shows the classification rate of the RBF-SVM model after the cross validation. The optimal RBF-SVM model was obtained when 8 PCs were used.

Table 4, shows the results of the RBF-SVM model of Bupleuri Radix. The classification rate was 96.6% in the

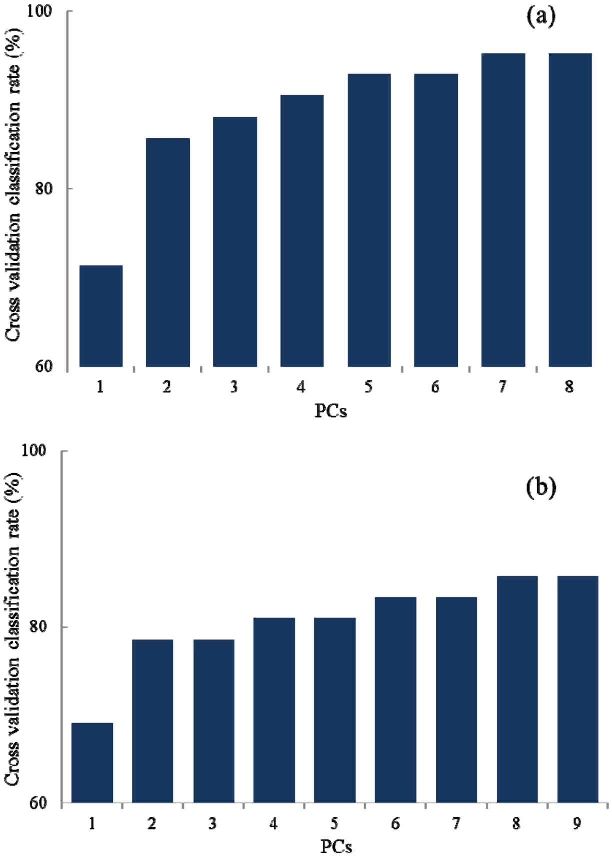


**Fig. 4.** Cross validation classification rates of QDA (a) and RBF-SVM (b) models at different PCs.

**Table 3.** Classification results by QDA model

| QDA | | Goheung | Whasun | Shanxi | Gansu | Shaanxi | Classification rate (%) | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | per group | all groups |
| Training set (n=89) | Goheung | 35 | 0 | 0 | 0 | 0 | 100 | 97.8 |
| | Whasun | 0 | 16 | 0 | 0 | 0 | 100 | |
| | Shanxi | 0 | 0 | 10 | 0 | 0 | 100 | |
| | Gansu | 0 | 0 | 0 | 17 | 0 | 100 | |
| | Shaanxi | 0 | 0 | 0 | 2 | 9 | 81.8 | |
| Test set (n=42) | Goheung | 17 | 0 | 0 | 0 | 0 | 100 | 95.2 |
| | Whasun | 0 | 7 | 0 | 0 | 0 | 100 | |
| | Shanxi | 0 | 0 | 5 | 0 | 0 | 100 | |
| | Gansu | 0 | 0 | 0 | 8 | 0 | 100 | |
| | Shaanxi | 0 | 0 | 0 | 2 | 3 | 60 | |

* The rows indicate the true sample class and that the columns refer to the observed class.

**Table 4.** Classification results by RBF-SVM model

| RBF-SVM | | Goheung | Whasun | Shanxi | Gansu | Shaanxi | Classification rate (%) | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | per group | all groups |
| Training set (n=89) | Goheung | 35 | 0 | 0 | 0 | 0 | 100 | 96.6 |
| | Whasun | 0 | 16 | 0 | 0 | 0 | 100 | |
| | Shanxi | 0 | 0 | 10 | 0 | 0 | 100 | |
| | Gansu | 0 | 0 | 0 | 17 | 0 | 100 | |
| | Shaanxi | 0 | 0 | 0 | 3 | 8 | 72.7 | |
| Test set (n=42) | Goheung | 17 | 0 | 0 | 0 | 0 | 100 | 85.7 |
| | Whasun | 0 | 7 | 0 | 0 | 0 | 100 | |
| | Shanxi | 0 | 0 | 3 | 0 | 2 | 60 | |
| | Gansu | 0 | 0 | 0 | 7 | 1 | 87.5 | |
| | Shaanxi | 0 | 0 | 1 | 2 | 2 | 40 | |

* The rows indicate the true sample class and that the columns refer to the observed class

training set and 85.7% in the test set, respectively. The RBF-SVM model did not show enhanced classification results than that of QDA model. According to a previous study,[29] the RBF-SVM model had a disadvantage when the dataset had many categories and relatively little training data for each category. In this study, the number of samples in the training set from Shanxi and Shaanxi was only about ten. It was not enough to obtain satisfactory results from the RBF-SVM model.

## Result and Discussion

In this study, near infrared spectroscopy combined with supervised pattern recognition was performed to classify Bupleuri Radix from five different habitats. Three supervised pattern recognition methods (PLS-DA, QDA and RBF-SVM) were compared to establish the classification model in this study. Among three classification models, the performance of the QDA model is improved than that of the others. When the number of PCs is 7, the classification rates of the QDA model for the training set and test sets were 97.8% and 95.2% respectively. The overall results demonstrate the feasibility of NIR spectroscopy combined with supervised pattern recognition for the geographical classification of Bupleuri Radix.

## Acknowledgement

This work is supported by the MOTIE (Ministry of Trade, Industry and Energy) R&D program (OSP grant 10039303) and partially supported by the Dongduk women University's sabbatical leave.

## References

(1) Ashour, M. L.; Wink, M. *J. Pharm. Pharmacol.* **2011**, *63*, 305-321.

(2) Bermejo Benito, P.; Abad Martínez, M. J.; Silván Sen, A. M.; Sanz Gómez, A.; Fernández Matellano, L.; Sánchez Contreras, S.; Díaz Lanza, A. M. *Life Sci.* **1998**, *63*, 1147-1156.

(3) Cheng, X. Q.; Li, H.; Yue, X. L.; Xie, J. Y.; Zhang, Y. Y.; Di, H. Y.; Chen, D. F. *J. Ethnopharmacol.* **2010**, *130*, 363-368.

(4) Sun, X. B.; Matsumoto, T.; Yamada, H. *J. Pharm. Pharmacol.* **1991**, *43*, 699-704.

(5) Wen, S.; Huifu, X.; Hao, H. *Immunopharmacol. Immunotoxicol.* **2011**, *33*, 433-437.

(6) Zhu, L.; Liang, Z. T.; Yi, T.; Ma, Y.; Zhao, Z. Z.; Guo, B. L.; Zhang, J. Y.; Chen, H. B. *BMC Complement. Altern. Med.* **2017**, *17*, 305-316.

(7) Li, X.; Jia, Y.; Song, A.; Chen, X.; Bi, K. *Yakugaku Zasshi* **2005**, *125*, 815-819.

(8) Bao, Y.; Li, C.; Shen, H.; Nan, F. *Anal. Chem.* **2004**, *76*, 4208-4216.

(9) Liau, B. C.; Hsiao, S. S.; Lee, M. R.; Jong, T. T.; Chiang, S. T. *J. Pharm. Biomed. Anal.* **2007**, *43*, 1174-1178.

(10) Lee, J.; Yang, D. H.; Suh, J. H.; Kim, U.; Eom, H. Y.; Kim, J.; Lee, M. Y.; Kim, J.; Han, S. B. *J. Chromatogr. B.* Analyt. Technol. Biomed. Life Sci. **2011**, *879*, 3887-3895.

(11) Huang, H. Q.; Su, J.; Zhang, X.; Shan, L.; Zhang, W. D. *J. Chromatogr. A.* **2011**, *1218*, 1131-1138.

(12) Tian, R. T.; Xie, P. S.; Liu, H. P. *J. Chromatogr. A.* **2009**, *1216*, 2150-2155.

(13) Qin, X.; Dai, Y.; Liu, N. Q.; Li, Z.; Liu, X.; Hu, J.; Choi, Y. H.; Verpoorte. R. *Planta Med.* **2012**, *78*, 926-933.

(14) Lin, X.; Xue, L.; Zhang, H.; Zhu, C. *Anal. Bioanal. Chem.* **2005**, *382*, 1610-1615.

(15) Gong, F.; Wang, B. T.; Chau, F. T.; Liang, Y. Z. *Anal. Lett.* **2005**, *38*, 2475-2492.

(16) McGoverin, C. M.; Weeranantanaphan, J.; Downey, G.; Manley, M. *J. Near Infrared Spec.* **2010**, *18*, 87-111.

(17) Chen, Y.; Xie, M. Y.; Yan, Y.; Zhu, S. B.; Nie, S. P.; Li, C.; Wang, Y. X.; Gong, X. F. *Anal. Chim. Acta* **2008**, *618*, 121-130.

(18) Luo, X. F.; Yu, X.;Wu, X. M.;Cheng, H. B.;Qu, H. B. *Microchem. J.***2008**, *90*, 8-12.

(19) Wang, L.; Lee, F. S. C.; Wang, X. *LWT-Food Sci. Technol.* **2007**, *40*, 83-88.

(20) Chen, Q.; Zhao, J.; Lin, H. *Spectrochim. Acta A. Mol. Biomol.*

*Spectrosc.* **2009**, *72*, 845-850.

(21) Lin, H.; Zhao, J.; Chen, Q.; Zhou, F.; Sun, L. *Spectrochim Acta. A. Mol. Biomol. Spectrosc.* **2011**, *79*, 1381-1385.

(22) Lee, D. Y.; Kim, S. H.; Kim, Y. C.; Kim, H. J.; Sung S. H. *Microchem. J.* **2011**, *99*, 213-217.

(23) Berrueta, L. A.; Alonso-Salces, R. M.; Héberger, K.; *J. Chromatogr. A.* **2007**, *1158*, 196-214.

(24) Li, B.; Wei, Y.; Duan, H.; Xi, L.; Wu, X. *Vib. Spectrosc.* **2012**, *62*, 17-22.

(25) Chiang, L. H.; Russell, E. L.; Braatz, R. D. *Chemometrics Intell. Lab. Syst.* **2000**, *50*, 243-252.

(26) Jiang, H.; Liu, G. H.; Xiao, X.; Yu, S.; Mei, C.; Ding, Y. *Food Anal. Methods* **2012**, *5*, 928-934.

(27) Luts, J.; Ojeda, F.; Van de Plas, R.; De Moor, B.; Van Huffel, S.; Suykens, J. A. *Anal. Chim. Acta* **2010**, *665*, 129-145

(28) Ballabio, D.; Consonni, V. *Anal. Methods* **2013**, *5*, 3790-3798.

(29) Rubin, T. N.; Chambers, A.; Smyth, P.; Steyvers, M. *Mach. Learn.* **2012**, *88*, 157-208.

Received December 18, 2017
Revised April 3, 2018
Accepted April 4, 2018