



## Short communication

## Simultaneous analysis of moisture, active component and cake structure of lyophilized powder for injection with diffuse reflectance FT-NIR chemometrics

Yang Li, Qi Fan\*, Sha Liu, Liqiong Wang

School of Pharmacy, Chongqing Medical University, Chongqing 400016, PR China

## ARTICLE INFO

## Article history:

Received 22 July 2010

Received in revised form

15 December 2010

Accepted 20 December 2010

Available online 30 December 2010

## Keywords:

Fourier transform near infrared

spectroscopy

Moisture

Active component

Cake structure

Lyophilized powder for injection

## ABSTRACT

This paper establishes a novel method for the simultaneous analysis of moisture, active component and cake structure of lyophilized powder for injection using diffuse reflectance Fourier transform near infrared (FT-NIR) chemometrics. The experiment indicated that the back-propagation artificial neural network (BP-ANN) was suitable for the content predictions of moisture and active component; the root mean square errors of prediction (RMSEPs) were 0.1471 and 0.0082, the correlation coefficients ( $R_s$ ) of prediction 0.9553 and 0.9891. And the self-organizing map (SOM) was adapted to the discrimination of cake structures; the prediction accuracy was 100.0%.

© 2010 Elsevier B.V. All rights reserved.

## 1. Introduction

The determinations of moisture and active component of lyophilized powder for injection are necessary because residual water may drastically affect the drug stability and the content of active component can indicate the drug effectivity [1]. On the other hand, the discrimination of lyophilized cake structures is also significant [2].

In this work, we discussed the feasibility of the simultaneous analysis of moisture, active component and cake structure of lyophilized powder for injection, such as lyophilized potassium sodium dehydroandrographolide succinate for injection, with diffuse reflectance FT-NIR chemometrics. And Karl–Fischer titration (KFT) [3], high performance liquid chromatography (HPLC) [4] and visual inspection (VI) [2] were respectively used as the reference methods. BP-ANN [5] and the partial least square (PLS) [5] were contrastively used for the determinations of moisture and active component, SOM [6] and the discriminant analysis (DA) [7] for the discrimination of cake structures. To our knowledge, there have not so far been any publications focused on it.

## 2. Experimental

## 2.1. Samples and materials

In the experiment, 115 samples of lyophilized powder for injection (80 mg/vial) and the reference substance (lot number: R0811066; purity: 97%, HPLC) of potassium sodium dehydroandrographolide succinate were prepared by Chongqing Yaoyou Pharmaceutical Co., Ltd (Chongqing, China). Methanol was HPLC grade, other chemicals were analytical grade and experimental water was purified water.

## 2.2. Instruments

The Fourier transform near infrared diffuse reflectance spectra (NIRDRS) were measured by Antaris II FT-NIR analyzer (Thermo Fisher Scientific, USA), furnished with an integrating sphere attachment and controlled with the software package RESULT 3.0. The reference values of moisture and active component were determined separately by DL31 Karl–Fischer titrator (Mettler Toledo, Switzerland) and LC-2010A high performance liquid chromatograph (Shimadzu, Japan).

## 2.3. Reference methods

## 2.3.1. Determination of moisture contents

The reference values of moisture contents were obtained by KFT. We divided the reference range of 85 samples, 0.68–2.66%, into

\* Corresponding author at: School of Pharmacy, Chongqing Medical University, Chongqing 400016, PR China. Tel.: +86 23 6848 5048; fax: +86 23 6848 5161.

E-mail address: [fanqi787@yahoo.com.cn](mailto:fanqi787@yahoo.com.cn) (Q. Fan).

five sub-ranges: 0.50–1.00%, 1.00–1.50%, 1.50–2.00%, 2.00–2.50% and 2.50–3.00%. The reference values of calibration set were distributed over the five sub-ranges, while the reference values of prediction set were distributed over the previous four sub-ranges. And the reference range of 65 calibration samples, 0.68–2.66%, covered that of 20 prediction samples, 0.82–2.46%.

### 2.3.2. Determination of the contents of active component

The reference values of the contents of active component, potassium sodium dehydroandrographolide succinate, were measured by HPLC. And the determination was performed on a Phenomenex Gemini C18 column (250 mm × 4.6 mm, 5 μm) with a mobile phase consisted of methanol-potassium dihydrogen phosphate (pH 3.0; 0.05 mol/L) (64:36, v/v) at a flow rate of 1.0 mL/min. The column temperature was maintained at 40 °C, the injection volume was 20 μL, and the detection was completed using the ultraviolet detector at 251 nm. We divided 30 samples into 25 calibration samples and 5 prediction samples using a procedure similar to that used in 2.3.1. The reference range of calibration set, 0.6518–0.8632 g/g, covered that of prediction set, 0.6666–0.8161 g/g.

### 2.3.3. Discrimination of cake structures

The cake structures of 40 samples were discriminated by VI in sealed cylindroid vials. The intact and defective cake subsets

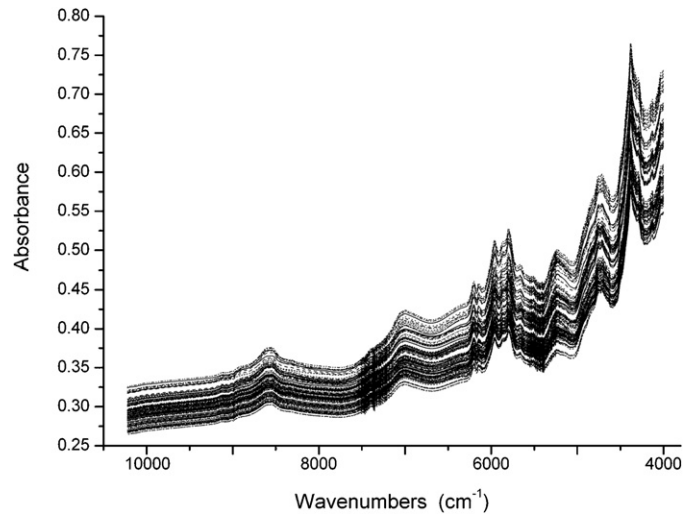


Fig. 1. Raw NIRDRS of lyophilized potassium sodium dehydroandrographolide succinate for injection.

were constituted respectively by 30 samples with the intact porous structure (20 for calibration, 10 for prediction) and 10 samples with the defective porous structure (7 for calibration, 3 for prediction). The calibration and prediction samples in two cake subsets were randomly selected.

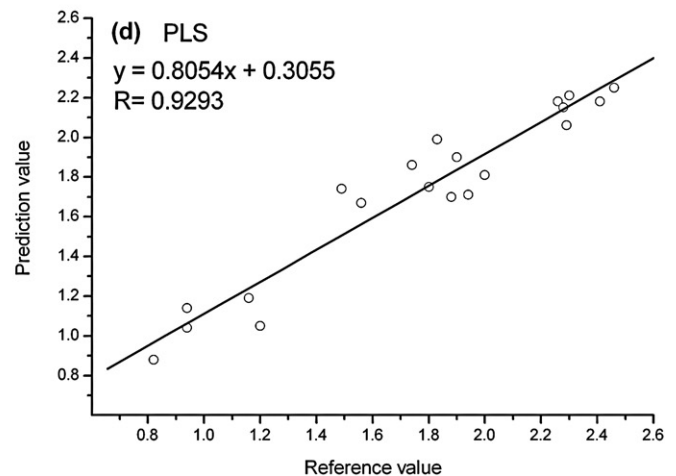
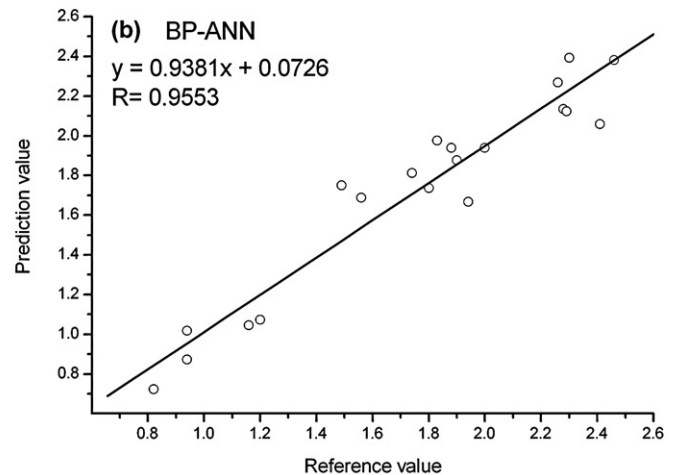
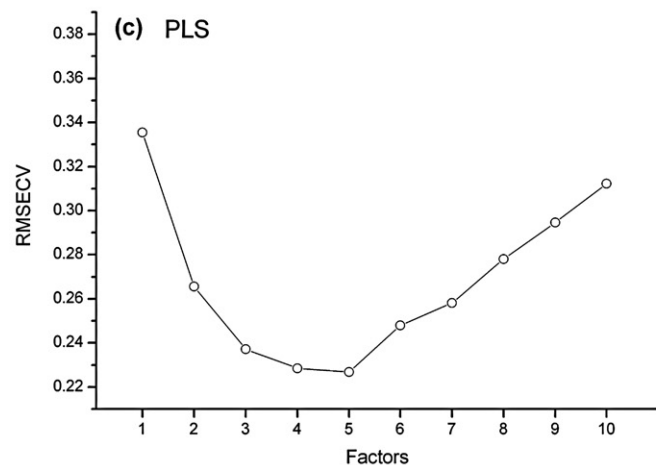
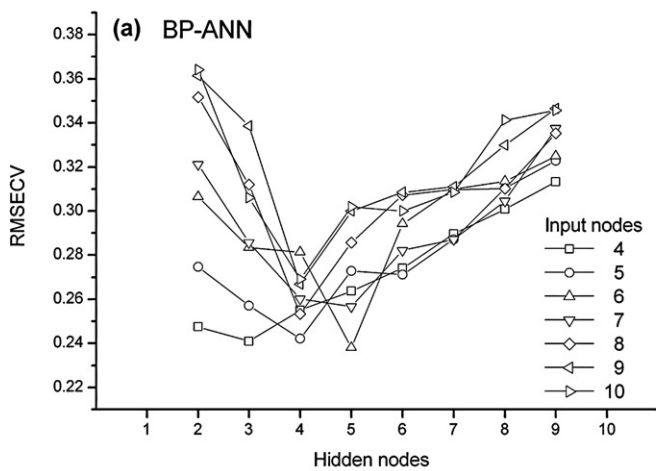


Fig. 2. Scatterplots for the moisture prediction models: (a) RMSECV values of the BP-ANN model; (b) prediction values and reference values of the prediction set for the BP-ANN model; (c) RMSECV values of the PLS model; (d) prediction values and reference values of the prediction set for the PLS model.

**Table 1**  
Optimal performances of the moisture prediction models using different preprocessing techniques.

Preprocessing techniques <sup>a</sup>	BP-ANN				PLS			
	PCs	RMSECV	RMSEP	R <sup>b</sup>	PCs	RMSECV	RMSEP	R <sup>b</sup>
No preprocess	4	0.2884	0.2031	0.8912	3	0.2826	0.2496	0.6882
MSC + first derivative	6	0.2687	0.1483	0.9427	1	0.2151	0.1628	0.9266
MSC + second derivative	8	0.2897	0.1921	0.8485	5	0.2267	0.1576	0.9293
SNV + first derivative	6	0.2376	0.1471	0.9553	1	0.2148	0.1639	0.9255
SNV + second derivative	8	0.2778	0.1908	0.9021	5	0.2266	0.1580	0.9293

<sup>a</sup> 7-point Savitzky–Golay smoothing and mean centering were included in each integration of preprocessing techniques.

<sup>b</sup> R was only for the prediction set.

## 2.4. NIR method

The NIRDRS of 115 samples were directly acquired in sealed cylindroid vials. Each spectrum was the average of 64 scans at 8 cm<sup>-1</sup> resolution over the range 10 000–4000 cm<sup>-1</sup> at 22 ± 1 °C and 55 ± 5% RH. The measurement background, the gold-plated inner wall of the integrating sphere, was also measured to automatically deduct the interferences of water vapor and CO<sub>2</sub> in air.

## 3. Results and discussion

The raw NIRDRS were shown in Fig. 1.

### 3.1. Prediction of moisture contents

The spectral sub-ranges, which were used to build the BP-ANN and PLS models, were automatically selected with the chemometric software TQ Analyst 8.0 (Thermo Fisher Scientific, USA). The regions were 8642–8217 cm<sup>-1</sup>, 7523–6206 cm<sup>-1</sup> and 5391–4948 cm<sup>-1</sup>.

To build the BP-ANN model, having low values of root mean square error of cross validation (RMSECV) and RMSEP, the NIRDRS were preprocessed by standard normal variate (SNV), 7-point Savitzky–Golay smoothing (5-, 7- and 9-point widths had been compared), first derivative and mean centering with TQ Analyst 8.0 (Table 1). The BP-ANN was built using three layers: an input layer which nodes represented the principle components (PCs) of the NIRDRS data for moisture content, an output layer with one node representing the prediction value of moisture content and one hidden layer [8]. The results were indicated in Fig. 2(a and b) and Table 1. Fig. 2(a) shows that the minimum RMSECV value was 0.2376 for 6 input nodes and 5 hidden nodes while the goal error between the training network output and the ideal output was defined as 0.05, that is, the optimal architecture of the BP-ANN was 6-5-1. Table 1 indicates that the BP-ANN had the low RMSEP value 0.1471 and the high R value 0.9553. And Fig. 2(b) displays the linear relation for the prediction set,  $y = 0.9381x + 0.0726$ . The BP-ANN algorithm was implemented by the mathematical modeling software Matlab 6.5 (The MathWorks, USA).

To build the PLS model, the NIRDRS were processed by multiplicative signal correction (MSC), 7-point Savitzky–Golay smoothing, second derivative and mean centering (Table 1). PCs of the NIRDRS data for moisture content were used as the factors for the PLS model. The results were indicated in Fig. 2(c and d) and Table 1. Fig. 2(c) describes that the minimum RMSECV value was 0.2267 as long as the number of factors was 5. In Table 1, the PLS had the low RMSEP value 0.1576 and the high R value 0.9293. And Fig. 2(d) demonstrates the linear relation for the prediction set,  $y = 0.8054x + 0.3055$ . The PLS algorithm was implemented using TQ Analyst 8.0.

The above discussion indicate that both of the BP-ANN and of the PLS models can be used to predict moisture contents of lyophilized powder for injection though the prediction performances of

the BP-ANN model were slightly better than those of the PLS model.

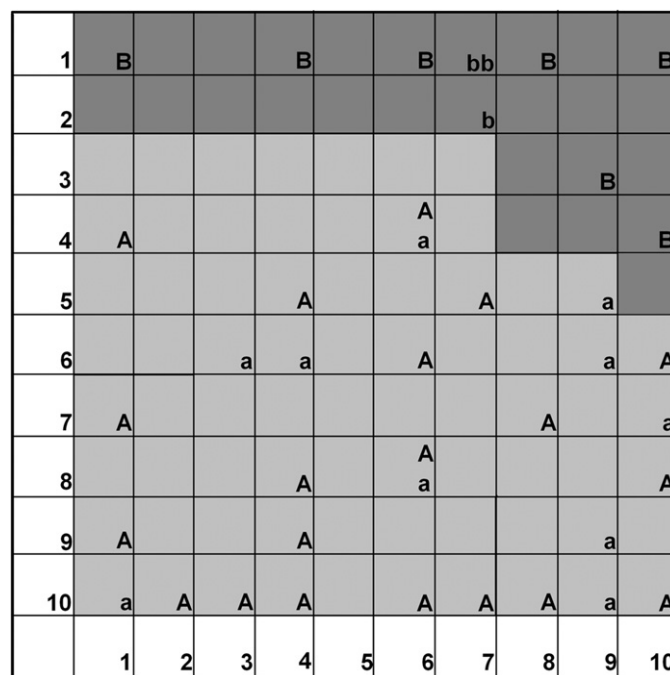
### 3.2. Prediction of the contents of active component

The contents of active component were predicted with the strategy similar to that used in 3.1. Three sub-ranges of the NIRDRS, 6995–6730 cm<sup>-1</sup>, 6283–5127 cm<sup>-1</sup> and 4873–4529 cm<sup>-1</sup>, were selected to build the BP-ANN and PLS models.

For the BP-ANN model, the used preprocessing techniques included MSC, 5-point Savitzky–Golay smoothing, first derivative and mean centering; the optimal architecture of the BP-ANN was 7-4-1 and the minimum RMSECV value was 0.0470 when the goal error was defined as 0.001; the RMSEP was 0.0082 and the R of prediction for the equation,  $y = 1.0366x - 0.0204$ , was 0.9891.

As for the PLS model, the used preprocessing techniques included MSC, 5-point Savitzky–Golay smoothing, second derivative and mean centering; the minimum RMSECV value was 0.0479 as long as the number of factors was 3; the RMSEP was 0.0242 and the R of prediction for the equation,  $y = 0.6788x + 0.2620$ , was 0.7964.

Consequently, it is different from the content prediction of moisture that the BP-ANN model is better than the PLS model for the content prediction of active component. The imaginable reason is



**Fig. 3.** Distribution map of the SOM (10 × 10) for cake structures: light gray blocks for intact cake subset (A, a) and dark gray blocks for defective cake subset (B, b); capitals (A, B) for calibration samples and lowercases (a, b) for prediction samples.

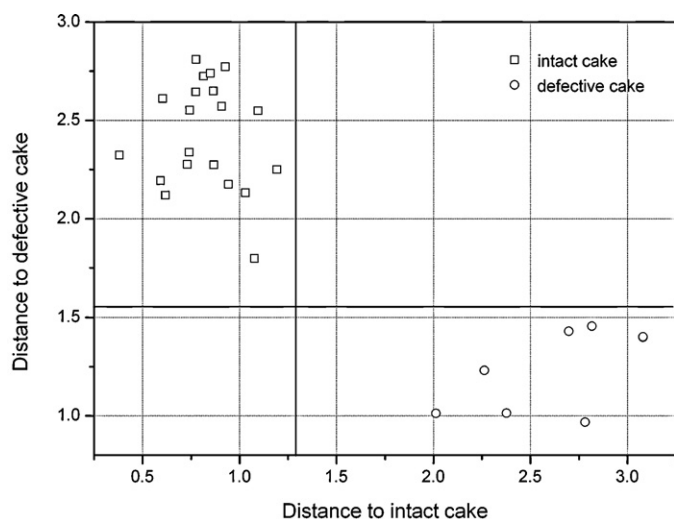


Fig. 4. Projection map of the DA of the calibration samples for cake structures.

that the NIR absorptions of active component are more complex and weaker than  $H_2O$ . BP-ANN, which is a non-linear method, may be more suitable to handle complex situation than PLS, which is a linear method, because it has stronger abilities of self-learning and self-adjust.

### 3.3. Prediction of cake structures

For the SOM and DA models, the full NIRDRS ( $1000\text{--}4000\text{ cm}^{-1}$ ) were selected automatically to build the models, processed by 7-point Savitzky–Golay smoothing, first derivative and mean centering, and reduced to 10 PCs which cumulative contribution rate was more than 97%, with TQ Analyst 8.0. The results were shown in Figs. 3 and 4.

In Fig. 3, the intact and defective cake subsets, not only calibration but also prediction samples, were completely separated into two areas without any errors. That is, the prediction accuracy of the SOM model was 100.0%. The SOM algorithm was implemented using Matlab 6.5.

Fig. 4 demonstrates that the calibration samples with the intact and defective porous structures were logically distributed with the

DA into two separate two-dimensional zones. The performance index of the DA model was 86.8%. The DA algorithm was done using TQ Analyst 8.0.

The results show that both of the SOM and of the DA models can be used to discriminate lyophilized cake structures although the prediction performances and the visualization function of the SOM model were better than those of the DA model.

## 4. Conclusion

The experiment has proved that moisture, active component and cake structure of lyophilized powder for injection can be simultaneously predicted by analyzing its NIRDRS with chemometrics: the BP-ANN models showed good performances for content predictions of moisture and active component; the SOM model appeared excellent prediction performances and visualization function for the discrimination of cake structures. Additionally, this dataset could be available for interested researchers to test new chemometric algorithms.

## Acknowledgments

We are grateful for the support provided by Chongqing Yaoyou Pharmaceutical Co., Ltd. and Thermo Fisher Scientific Co., Ltd.

## References

- [1] S.P. Schwendeman, H.R. Costantino, R.K. Gupta, G.R. Siber, A.M. Klivanov, R. Langer, Stabilization of tetanus and diphtheria toxoids against moisture-induced aggregation, *Proc. Natl. Acad. Sci. U.S.A.* 92 (1995) 11234–11238.
- [2] W. Abdelwahed, G. Degobert, S. Stainmesse, H. Fessi, Freeze-drying of nanoparticles: formulation, process and storage considerations, *Adv. Drug Deliv. Rev.* 58 (2006) 1688–1713.
- [3] Food and Drug Administration, Guideline for the determination of residual moisture in dried biological products (Docket No. 89D-0140).
- [4] Y. Ma, H. Ding, Determination of potassium sodium 14-deoxy-11,12-didehydroandrographolide succinate by HPLC, *Chin. J. Pharm.* 37 (2006) 123–124.
- [5] M. Otto, *Chemometrics*, Science Press, Beijing, 2003.
- [6] T. Kohonen, *Self-organizing Maps*, third ed., Springer, Berlin, 2001.
- [7] P.C. Jurs, G.A. Bakken, H.E. McClelland, Computational methods for the analysis of chemical sensor array data from volatile analytes, *Chem. Rev.* 100 (2000) 2649–2678.
- [8] Y. Roggo, P. Chalus, L. Maurer, C. Lema-Martinez, A. Edmond, N. Jent, A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies, *J. Pharm. Biomed. Anal.* 44 (2007) 683–700.