

## ***In-situ* monitoring of saccharides removal of alcohol precipitation using near-infrared spectroscopy**

Hongxia Huang\* and Haibin Qu<sup>†,‡</sup>

\*School of Medicine, Zhejiang University City College  
Hangzhou, Zhejiang 310015, P. R. China

<sup>†</sup>Pharmaceutical Informatics Institute, Zhejiang University  
Hangzhou, Zhejiang 310058, P. R. China

<sup>‡</sup>quhb@zju.edu.cn

Received 24 April 2018

Accepted 1 July 2018

Published 20 July 2018

As unsafe components in herbal medicine (HM), saccharides can affect not only the drug appearance and stabilization, but also the drug efficacy and safety. The present study focuses on the in-line monitoring of batch alcohol precipitation processes for saccharide removal using near-infrared (NIR) spectroscopy. NIR spectra in the 4000–10,000-cm<sup>-1</sup> wavelength range are acquired *in situ* using a transfectance probe. These directly acquired spectra allow characterization of the dynamic variation tendency of saccharides during alcohol precipitation. Calibration models based on partial least squares (PLS) regression have been developed for the three saccharide impurities, namely glucose, fructose, and sucrose. Model errors are estimated as the root-mean-square errors of cross-validation (RMSECVs) of internal validation and root-mean-square errors of prediction (RMSEPs) of external validation. The RMSECV values of glucose, fructose, and sucrose were 1.150, 1.535, and 3.067 mg·mL<sup>-1</sup>, and the RMSEP values were 0.711, 1.547, and 3.740 mg·mL<sup>-1</sup>, respectively. The correlation coefficients (*r*) between the NIR predictive and the reference measurement values were all above 0.94. Furthermore, NIR predictions based on the constructed models improved our understanding of sugar removal and helped develop a control strategy for alcohol precipitation. The results demonstrate that, as an alternative process analytical technology (PAT) tool for monitoring batch alcohol precipitation processes, NIR spectroscopy is advantageous for both efficient determination of quality characteristics (fast, *in situ*, and requiring no toxic reagents) and process stability, and evaluating the repeatability.

**Keywords:** Herbal medicine; alcohol precipitation; near-infrared spectroscopy; saccharides removal; process analytical technology.

<sup>‡</sup>Corresponding author.

This is an Open Access article published by World Scientific Publishing Company. It is distributed under the terms of the Creative Commons Attribution 4.0 (CC-BY) License. Further distribution of this work is permitted, provided the original work is properly cited.

## 1. Introduction

The safety and efficacy of herbal medicine (HM) are dependent not only on the quality of medicinal materials but also on the manufacturing processes. Owing to the complexity of HM, a number of purification technologies have been used by the pharmaceutical industry to manufacture HMs with high safety standards and efficacy. Efficacy ingredients, toxic and harmful components, and potential unsafe components are considered the three main groups for analysis in the purification processes.<sup>1</sup> The removal of potential unsafe components can improve drug safety and reduce drug dosage. Although not considered as efficacy components, saccharides, such as glucose, fructose, and sucrose, are commonly found in HM extracts. These saccharides in HMs can affect not only the drug appearance and stability, but also the efficacy and safety. Color, degradation of glucose, fructose, and sucrose, and organic acid formation have been shown to be highly correlated.<sup>2</sup> Coca *et al.*<sup>3</sup> further confirmed that colored compounds were formed in juices and syrups due to sugar degradation and reactions between amine compounds and carbohydrates. Researchers have also focused on the transformation pathways and products of these saccharides. Yang and Montgomery<sup>4</sup> and Asghari and Yoshida<sup>5</sup> reported that fructose and other saccharides can possibly produce 5-hydroxymethylfurfural (5-HMF) under certain conditions. The production of 5-HMF, which has reported cytotoxicity, genotoxicity, and tumorigenicity, is limited at different safe concentrations in different pharmacopeia.<sup>6</sup> Therefore, the further reduction and removal of saccharides during manufacturing process is still needed to assure high quality of their end products.

Alcohol precipitation is among the most effective methods for saccharide removal, with the advantage of being simple, rapid, easily scalable, and cost-effective.<sup>7</sup> This method has also been identified as a key operation affecting the downstream processes and even product quality, based on previous research and risk assessment conducted by ICH-Q8 and ICH-Q9 guidelines.<sup>8–10</sup> Traditionally, various laboratory-based analytical techniques, such as colorimetric measurements, high-performance liquid chromatography (HPLC), and capillary electrophoresis, have been used for the quantitative determination of saccharides in HM, biological systems, fruit juice, and wine.<sup>11–13</sup> However, these

offline methods suffer from various disadvantages, such as time delays, nonrepresentative sample collection, analytical errors in sample preparation, and labor-intensive analysis.<sup>9,14</sup>

In contrast, process analytical technology (PAT) tools initiated by the Food and Drug Administration (FDA)<sup>15</sup> may overcome the difficulties of offline laboratory techniques and allow at-line, in-line, or online measurements of critical process performance and quality attributes. Following such initiative, various spectroscopic (UV/vis, IR, and Raman) and chromatographic methods for process monitoring have been developed. Among these methods, near-infrared (NIR) spectroscopy is the most widely used in the pharmaceutical industry,<sup>16–19</sup> mainly owing to its rapid turnaround (in seconds or minutes), nondestructive nature, high sensitivity, and minimal or no sample preparation.<sup>20</sup> NIR spectroscopy, a type of vibration spectroscopy, has a wavelength range of 700–2500 nm (14,300–4000 cm<sup>-1</sup>). Fundamental absorptions found in mid-IR are mainly due to overtones and combination bands of hydrogen bonds, such as –OH, –NH, and –CH.<sup>21</sup> Luypaert *et al.*<sup>22</sup> published a review on pharmaceutical applications of NIR, including NIR technique development in the pharmaceutical field, where it could be applied to raw material identification, manufacturing process monitoring, and final product release. A book reviewing the pharmaceutical and medical applications of NIR spectroscopy by Ciurczak and Drennen<sup>17</sup> describes NIR applications at different stages of manufacture, including blending, granulation, drying, and coating. Compared with conventional laboratory analysis, a significant amount of analysis time can be saved and sampling errors reduced by *in-situ* processes measurement.<sup>9</sup> However, the application of NIR spectroscopy to HM manufacturing processes remains limited. HM manufacture is complex and involves many process stages and constituents, presenting a great challenge for in-line monitoring and quality control. NIR spectroscopy can be used to measure many chemical compositions and physical properties (such as temperature, viscosity, and granularity) and can, therefore, play an important role in HM quality control and improving the continuous process control strategies, quality verification, and real-time release (RTR).<sup>23–25</sup> However, NIR calibrations often require a sampling procedure for reference measurements, which is particularly important in PAT applications using spectroscopic

probes, because reference values of samples are biased against the real process state.

Herein, the alcohol precipitation process of Danshen (*Salvia miltiorrhiza bunge*) was selected for study. Danshen preparations made from aqueous extracts of *Salvia miltiorrhiza* root are among the most widely used HMs for the treatment of cardiovascular disease, heart stroke, and cerebrovascular disease.<sup>26–28</sup> Potential unsafe saccharide components, such as glucose, fructose, sucrose, and raffinose, are commonly found in Danshen extracts.<sup>1</sup> Herein, NIR spectroscopy was evaluated as a method for monitoring the *in-situ* batch alcohol precipitation processes for saccharide removal. An NIR transmittance probe was directly inserted into an alcohol precipitation system (*in situ*) and corresponding spectral measurements were obtained in several seconds. The spectral results were subsequently used to model three key quality attributes: glucose, fructose, and sucrose. These attributes were selected because they are nonbioactive impurities and are typically used to characterize the effectiveness of batch alcohol precipitation processes. NIR predictions were then employed directly to evaluate the batch processes and enhance the process understanding. All experiments were performed on a laboratory scale to simulate the industrial process of alcohol precipitation for saccharide removal from Danshen.

## 2. Experiment and Methods

### 2.1. Materials

Danshen extracts and 95% EtOH were obtained from Chiatai Qingchunbao Pharmaceutical Co., Ltd. (Hangzhou, China) and Changqin Chemical Co., Ltd. (Hangzhou, China), respectively. Standards for D-glucose, D-fructose, and sucrose were purchased from Sinopharm Chemical Reagent Co., Ltd. (Shanghai, China). HPLC-grade acetonitrile and formic acid were supplied by Merck (Darmstadt, Germany) and Tedia Company (Fairfield, OH, USA), respectively. Deionized water was prepared using a Milli-Q water purification system (Milford, MA, USA).

### 2.2. Experiment setup and design

The batch alcohol precipitation process was performed in a custom glass beaker with a maximum

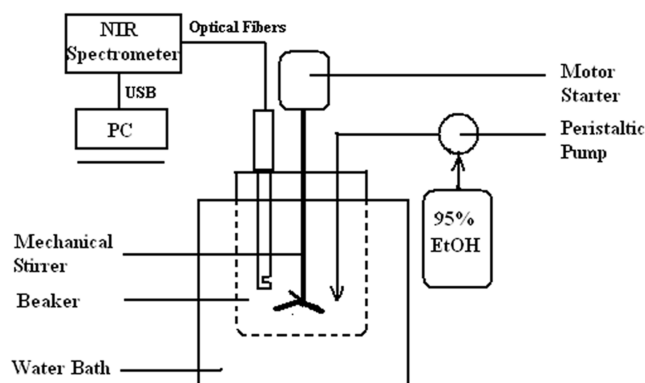


Fig. 1. Schematic of experiment setup for alcohol precipitation of Danshen.

volume of about 4 L. The experimental setup is shown in Fig. 1. An aqueous solution of Danshen (approximately 1.23 kg) was introduced into the beaker and kept in a water bath at 20°C under mechanical stirring with a three-impeller stirrer and a motor starter. Stirring was continued while 95% EtOH was added with a peristaltic pump. The rate of EtOH addition was set at 100 mL/min. The solution was then allowed to stand for an hour without stirring. The NIR transmittance probe was then carefully inserted into the beaker at a fixed point to collect the NIR spectra continually during alcohol precipitation.

All experiments were conducted on a laboratory scale. A design matrix with two factors, six center-point (CP) replicates, and two abnormal experimental runs was constructed. Two experimental factors, EtOH volume and initial density (solid concentration of the aqueous solution of Danshen), were included in the design of experiments (DoE), as summarized in Table 1. In addition to DoE, abnormal operation conditions were added to two batches (Table 2) named process deviation (PD) batches to simulate the real-world process deviations and assess model capability.

### 2.3. NIR instrumentation

NIR spectra were continuously collected *in situ* during alcohol precipitation using a Thermo Scientific<sup>TM</sup> Antaris II Fourier Transform (FT) NIR analyzer equipped with a transmittance probe with an adjustable optical path length (S-650, Thermo Nicolet Corporation, WI, USA). The transmittance probe (path length fixed to about 2 mm) was connected directly to the NIR spectrometer via an optic

Table 1. Factor settings for the design of experiments.

Batch name (abbreviation)	Initial density ( $\text{kg} \cdot \text{m}^{-3} 10^3$ )	EtOH volume (L)
Center point 1 (CP1)	1.23	3.0
Center point 2 (CP2)	1.23	3.0
Center point 3 (CP3)	1.23	3.0
Center point 4 (CP4)	1.23	3.0
Center point 5 (CP5)	1.23	3.0
Center point 6 (CP6)	1.23	3.0
Experiment 7 (DoE7)	1.23	2.0
Experiment 8 (DoE8)	1.18	3.0
Experiment 9 (DoE9)	1.23	4.0
Experiment 10 (DoE10)	1.28	3.0
Process deviation 11 (PD11)	1.23	3.0
Process deviation 12 (PD12)	1.23	3.0

Table 2. Process disturbances introduced to batches PD10 and PD11.

Batch name (abbreviation)	Description of process deviation
Process deviation (PD11)	Agitation failure: Stirrer stopped at $t = 16$ min, restarted at $t = 26$ min
Process deviation (PD12)	Peristaltic pump failure: Increase rate of EtOH addition to 200 mL/min at $t = 11$ min; Decrease rate of EtOH addition to 100 mL/min at $t = 22$ min; Decrease rate of EtOH addition to 50 mL/min at $t = 32$ min

fiber and inserted into the beaker to record the process spectra. Each spectrum was collected in the 5000–10,000- $\text{cm}^{-1}$  range with a total of 64 scans at a resolution of 8.0  $\text{cm}^{-1}$ . The software was programmed to acquire one spectrum every minute. The total duration of the batch process was 80 min, comprising 10 min for raw material addition, 37 min for 95% EtOH addition, and 33 min for the standing phase.

## 2.4. Samples

To obtain a good sampling distribution, samples (2.0 mL) were taken at random every 5–30 min for each batch. A total of 115 samples were collected

throughout alcohol precipitation from different experimental batches, and their glucose, fructose, and sucrose contents were analyzed by HPLC using reference methods (see Sec. 2.5 for the specific method). To obtain representative samples, both CP batches and DoE batches with a wide range of saccharides were used during model development. Glucose, fructose, and sucrose models used in this study were developed based on 100 samples from these CP and DoE batches. A total of 15 samples from the PD batches (eight samples from PD11 and seven samples from PD12) were used as an external validation set.

## 2.5. Reference methods

The glucose, fructose, and sucrose contents were determined by HPLC-evaporative light scattering detector (ELSD) analysis. An Agilent 1100 series HPLC instrument (Agilent Technologies, Waldbronn, Germany) equipped with a vacuum degasser, a quaternary gradient pump, an autosampler, and a thermostated column compartment was used. The signal from a Sedex75 ELSD (Sedere, France) was transmitted to the Chemstation for processing by an Agilent 35900E A/D interface (Agilent Technologies, Santa Clara, USA). Chromatographic separations were performed in a WATO44355 carbohydrate column (250 mm  $\times$  4.6 mm, 4  $\mu\text{m}$ , Waters, MA) at 1.0  $\text{mL} \cdot \text{min}^{-1}$  using acetonitrile–water (78:22, v/v) as the mobile phase. The column temperature was set at 30  $^{\circ}\text{C}$ , and the injection volume was 10  $\mu\text{L}$ . Samples of crude extract and alcohol precipitation liquid were diluted twofold and 20-fold, respectively, and centrifuged at 12,000 rpm for 10 min using a centrifuge (Eppendorf Co., Germany) before HPLC-ELSD analysis.

## 2.6. Chemometrics

Calibration models based on NIR spectra for glucose, fructose, and sucrose were developed using the partial least squares (PLS) algorithm.<sup>29</sup> PLS regression analysis was performed to establish a linear relationship between the two matrices, NIR spectral data ( $\mathbf{X}_{n \times m}$ ) and saccharide concentration ( $\mathbf{Y}_{n \times k}$ ), which were recorded for  $n$  samples. The objective of PLS was to project the data down onto a number of latent variables (LVs), such as  $t_i$  and

$\mathbf{u}_i$ , where  $i$  is the number of latent variables, and then develop a regression model between  $\mathbf{t}_i$  and  $\mathbf{u}_i$ :

$$\mathbf{u}_i = b_i \mathbf{t}_i + \mathbf{e}_i, \quad (1)$$

where  $\mathbf{e}_i$  is a vector of errors and  $b_i$  is the regression coefficient of the latent variables  $\mathbf{X}$  and  $\mathbf{Y}$  estimated by  $b_i = \mathbf{t}_i^T \mathbf{u}_i / (\mathbf{t}_i^T \mathbf{t}_i)$ :

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E} = \sum_{i=1}^k \mathbf{t}_i \mathbf{p}_i^T + \mathbf{E}, \quad (2)$$

$$\mathbf{Y} = \mathbf{UQ}^T + \mathbf{F} = \sum_{i=1}^k \mathbf{u}_i \mathbf{q}_i^T + \mathbf{F}, \quad (3)$$

where  $\mathbf{T}$  and  $\mathbf{P}$  are the scores and loadings (or the contributions) for  $\mathbf{X}$ ,  $\mathbf{U}$  and  $\mathbf{Q}$  are the scores and loadings for  $\mathbf{Y}$ .  $\mathbf{E}$  and  $\mathbf{F}$  are the residuals of  $\mathbf{X}$  and  $\mathbf{Y}$  after extracting the first  $i$  pairs of latent variables.

PLS regression is a multivariate technique with no restriction on the number of wavelengths, and is especially useful for ranking deficiency, when the number of samples is less than that of variations or in the case of highly correlated variables.<sup>30</sup> Further details on PLS modeling can be found in the literature.<sup>31</sup> The two most commonly used validation methods are cross-validation and test set validation. The  $k$ -fold cross-validation is a commonly used cross-validation technique for estimating the prediction error. From  $k$  subsamples, a single subsample is retained as the validation data to test the model, and the remaining  $k - 1$  subsamples are used as training data. The cross-validation process is then repeated  $k$  times (the folds), with each of the  $k$  subsamples being used exactly once as the validation data. The  $k$  results from the folds can then be averaged (or otherwise combined) to produce a single estimate. This method has the advantage that all observations are used for both training and validation, and each observation is used for validation exactly once. In this study, tenfold cross-validation was applied as the internal validation. All data were chemometrically processed using Unscrambler 9.6 (Camo, Oslo, Norway).

The performance of the constructed model was evaluated using the correlation coefficient ( $r$ ), the root-mean-square error of cross-validation (RMSECV), and the root-mean-square error of prediction (RMSEP). The  $r$  value between the NIR prediction and measured value was calculated as

follows:

$$r = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}, \quad (4)$$

where  $y_i$  and  $\hat{y}_i$  are measured value and predicted value for the  $i$ th sample, respectively, and  $\bar{y}$  is the mean of the reference results for all samples in the training and test sets.

RMSECV was determined from the cross-validation, calculated as follows:

$$\text{RMSECV} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_{\text{CV}_i} - y_i)^2}{n}}, \quad (5)$$

where  $n$  is the number of calibration samples,  $\hat{y}_{\text{CV}_i}$  is the cross-validation estimate of sample  $i$ . The optimum number of LVs was determined by plotting RMSECV against the rank (number of factors used in the model). The minimum indicates the optimum rank or number of factors that minimize the model prediction error.

RMSEP was used as a model accuracy indicator of the external testing, and was calculated as follows:

$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^m (\hat{y}_i - y_i)^2}{m}}, \quad (6)$$

where  $m$  is the number of samples in the test set.

Different spectra preprocessing techniques,<sup>32</sup> such as derivatives, standard normal variate (SNV), and multiplicative signal correction (MSC), were applied to minimize the physical differences of samples, especially in terms of particle size and undesirable systematic variations.<sup>33</sup> In this study, considering the characteristics of the spectra, SNV was applied for preprocessing the spectra to minimize the effect due to different kinds of interference. This method centers each spectrum at around zero by subtracting from the average (additive adjustment) or dividing by the standard deviation (multiplicative adjustment) in all variables of the spectrum.

## 2.7. In-line monitoring advantages

Predictive saccharide values based on the constructed models were produced every minute to provide updates on the process status. In-line operation and integration with the predictive system reduced the sampling and test requirements, and, therefore, freed up resources. An increase in sampling frequency led to more effective monitoring and

control of alcohol precipitation for saccharide removal. Process deviations and disturbance were detected in a timely fashion based on the process trajectory and evolution of batches for saccharide removal performance. The process trajectory of alcohol precipitation was determined using SIMCA-P+ 12.0.1 software (Umetrics, Umeå, Sweden).

### 3. Results and Discussion

#### 3.1. NIR process spectra

NIR spectra of the beaker contents were measured every minute (*in-situ* measurement). Figure 2 shows the raw and preprocessed spectra of a normal-operating-conditions batch run of alcohol precipitation from Danshen using SNV. Preprocessing enabled most spectral variations to be eliminated due to physical differences. The processed NIR spectra exhibited intense absorbance bands in the

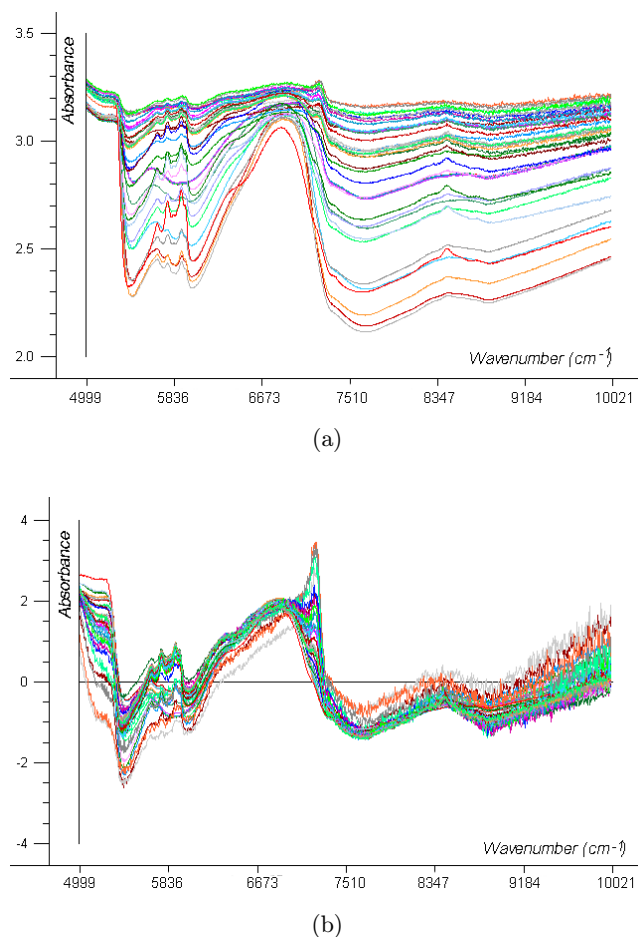


Fig. 2. NIR spectra obtained every minute from *in-situ* monitoring during the alcohol precipitation: (a) raw NIR spectra and (b) preprocessed spectra with SNV.

5300–7274- $\text{cm}^{-1}$  wavelength region. The peak at around 5361  $\text{cm}^{-1}$  might be attributed to the absorption of a combination of O–H stretching and bending, while the peak around 6888  $\text{cm}^{-1}$  was likely due to the first overtone of O–H stretching in alcohol and water.<sup>34</sup> The 5682–6024- $\text{cm}^{-1}$  region was related to the first overtone of C–H stretching modes in  $\text{CH}_3$  and  $\text{CH}_2$  groups.<sup>35</sup>

#### 3.2. Saccharides calibration

The variability observed among the 100 Danshen samples used to model the glucose, fructose, and sucrose parameters could be measured by dividing the standard deviation of each parameter by the corresponding range (Table 3). Three PLS1 models, one each for glucose, fructose, and sucrose, were developed based on NIR spectra acquired *in situ*. For all calibration models, the best preprocessing method was found to be SNV and the optimal region was 5681–6024  $\text{cm}^{-1}$  combined with 6329–7274  $\text{cm}^{-1}$ . Optimal PLS models of glucose, fructose, and sucrose, cross-validated with 100 samples, were obtained using four components, and the corresponding RMSECV values were 1.150, 1.535, and 3.067  $\text{mg} \cdot \text{mL}^{-1}$ , respectively.

Figure 3 represents the reference and cross-validated PLS predicted values for glucose, fructose, and sucrose. Good agreement was found between reference and predictive values both in high correlation coefficient and low RMSECV. The models exhibited a good utility value for the quality control of alcohol precipitation. The RMSEP values of external validation for glucose, fructose, and sucrose (Table 4) were 0.711, 1.547, and 3.740  $\text{mg} \cdot \text{mL}^{-1}$ , respectively, which were obtained in line. As shown in Table 4, higher sugar concentrations resulted in larger predictive errors. A possible explanation was that sampling at high sugar concentrations presented a significant challenge owing to the rapid decrease in sugar content at the beginning of the purification process with increasing ethanol mass fraction.

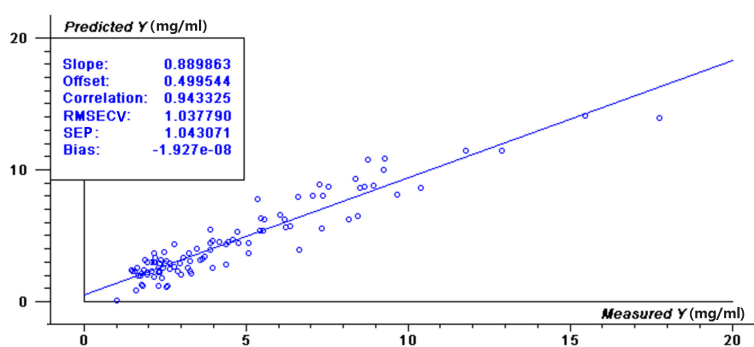
#### 3.3. Implementation of in-line monitoring

The effectiveness of NIR spectroscopy in predicting the tendency of these saccharide parameters during the batch alcohol precipitation process could enhance the process understanding. Figure 4(a)

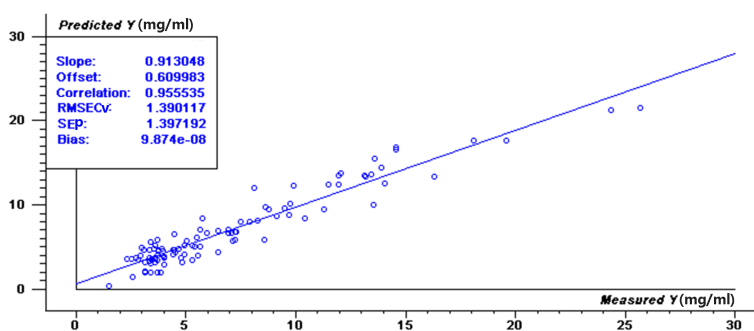
Table 3. Concentration ranges of the samples in the training set (a) and external testing set (b).

Sample set	Unit	Range	Sample number	Mean	S.D. <sup>a</sup>
(a)			100		
Glucose	mg · mL <sup>-1</sup>	1.02–17.75		4.56	3.14
Fructose	mg · mL <sup>-1</sup>	1.47–25.71		7.09	4.77
Sucrose	mg · mL <sup>-1</sup>	2.41–49.98		12.21	8.43
(b)			15		
Glucose	mg · mL <sup>-1</sup>	1.73–13.05		3.40	2.94
Fructose	mg · mL <sup>-1</sup>	2.86–22.41		5.35	5.04
Sucrose	mg · mL <sup>-1</sup>	4.67–16.91		7.47	3.82

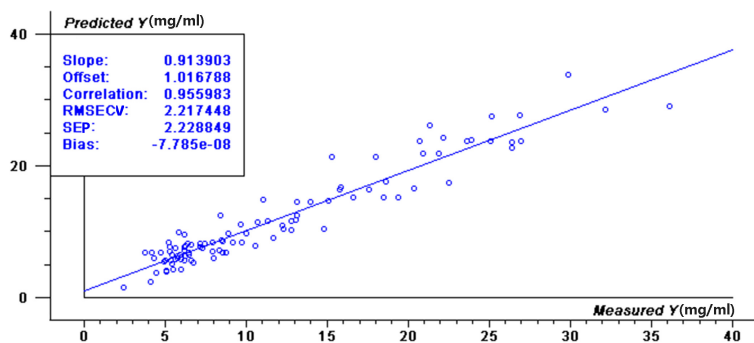
Note: <sup>a</sup>S.D. = standard deviation.



(a)



(b)



(c)

Fig. 3. PLS NIR-model validation plots for saccharides of Danshen: (a) glucose, (b) fructose, and (c) sucrose.

Table 4. Predictive/reference values of external validation for glucose, fructose, and sucrose models.

Sample	Glucose ( $\text{mg} \cdot \text{mL}^{-1}$ )		Fructose ( $\text{mg} \cdot \text{mL}^{-1}$ )		Sucrose ( $\text{mg} \cdot \text{mL}^{-1}$ )	
	Predictive	Reference	Predictive	Reference	Predictive	Reference
1	11.156	13.052	17.177	17.177	28.048	14.030
2	4.162	4.310	6.250	6.250	10.928	11.574
3	2.502	3.212	3.984	3.984	7.369	8.060
4	2.533	2.385	3.924	3.924	6.704	6.025
5	1.693	2.031	2.714	2.714	4.887	5.418
6	1.27	1.731	2.133	2.133	4.498	4.757
7	1.743	1.854	2.843	2.843	5.01	4.673
8	1.743	1.854	2.843	2.843	5.01	4.673
9	7.074	6.224	10.932	10.932	19.102	16.912
10	3.976	3.658	6.027	6.027	10.137	8.991
11	2.196	2.687	3.504	3.504	6.349	6.480
12	1.442	2.103	2.401	2.401	4.186	5.288
13	0.964	2.139	1.694	1.694	3.488	5.284
14	1.44	1.885	2.411	2.411	4.432	5.079
15	1.888	1.851	3.076	3.076	5.44	4.839
RMSEP	0.711		1.547		3.740	

shows the results of sucrose PLS calibration during the dynamic alcohol precipitation process generated using data from one CP batch, as compared with the reference measurements. Good agreement between predictive and reference measurement values was observed, except for at the timepoint 40. At timepoints 37–47 during the end stage of alcohol addition, precipitates from adsorption cross-linked with each other and agglomerated to form larger precipitates. NIR spectra showed interference due to these particle variations, which made prediction challenging. Six CP batches with sucrose variation

were predicted using the sucrose PLS calibration shown in Fig. 4(b), in which three stages (feeding at timepoints 1–10, EtOH addition stage at timepoints 11–47, and standing at timepoints 48–80) of the batch process are distinguished. The predictive sucrose values remained constant until EtOH addition at the filling stage. The sucrose content tended to decrease during the purification process, probably owing to the solubility of sucrose being significantly reduced with an increasing ethanol mass fraction. The number of small particles increased and became flocculent, but did not aggregate until the

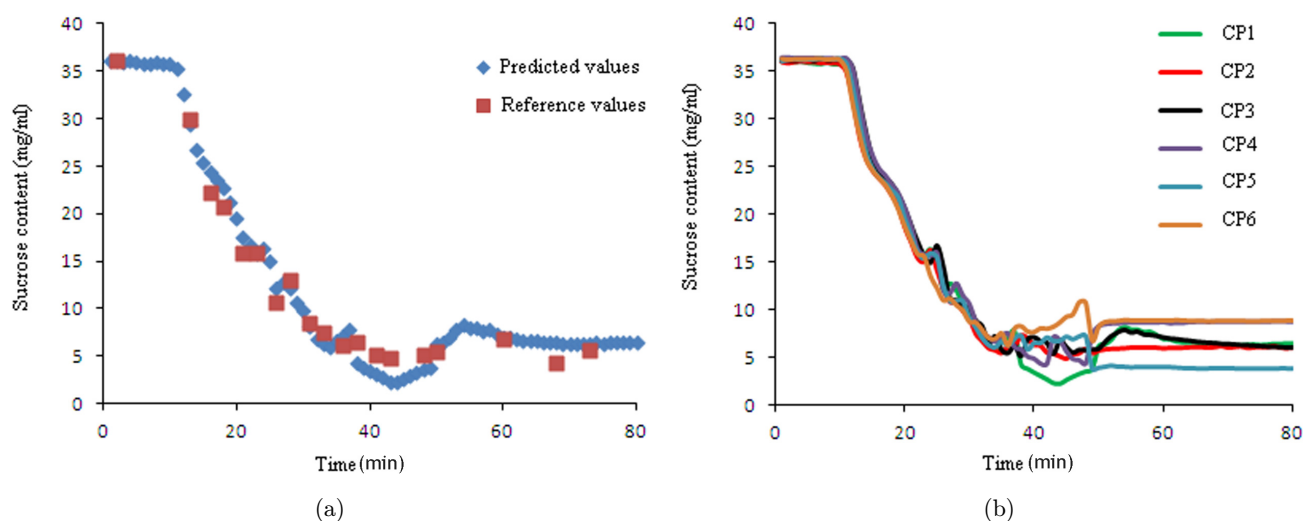


Fig. 4. PLS models for in-line monitoring applications: (a) sucrose prediction during the alcohol precipitation for certain batch runs and (b) process trajectory and evolution of batches for sugar removal based on the six CP batches.



purification had ended. When the alcohol content of the supernatant reached a certain level at timepoints 37–47 (the critical point), precipitates of small particles continued to cross-link into flocs, followed by aggregate formation. During this process, various particles increased the complexity of the system, making the process trajectory highly volatile. The sucrose content in the supernatant after about 60 min was stable, implying that the system had reached a balance. The other two kinds of saccharides (not shown here) exhibited the same trends and behaviors. Furthermore, the results in the chart highlighted the reproducibility of the CP batches based on in-line NIR measurement of saccharides.

In the derived principle component analysis (PCA) model, the first principle component (PC) captured 99.9% of the overall variation, representing a good summary of batch variations. A control chart of the three saccharides based on the six CP batches is shown in Fig. 5(a). The control limits of the batch control chart corresponded to an average

saccharide content of  $\pm 3$  standard deviations. Two PD batches were used for validation, with both abnormal batches exceeding the control limits at certain timepoints. Agitator failure, caused by the agitator stirring stopping at timepoint 16, was introduced to create process deviation during the alcohol precipitation of batch PD11. As a result, batch PD11 evolved outside the control limits accurately at timepoint 16. The process trajectory fell back into the control line until the paddle returned to normal at timepoint 26. Similar behavior was observed in the batch PD12 with a peristaltic pump failure. Contribution plots in Figs. 5(b) and 5(c) were used to determine which of the original variables were most related with these phenomena. Three saccharides were positively correlated during the alcohol precipitation process. The trajectories of the PD batches recovered and returned to within normal limits after the process disturbance disappeared. The agitation break or peristaltic pump failure for a short period was shown to have an insignificant effect on the end product, which was in

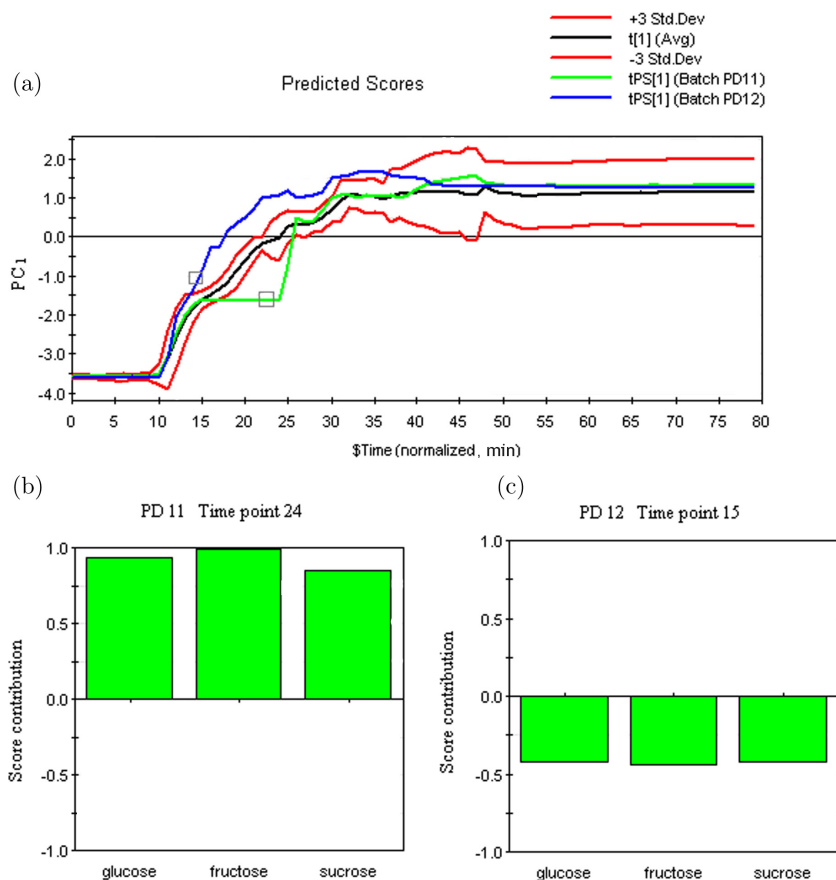


Fig. 5. Process monitoring of the PD baths: (a) A control plot of the PD baths based on the six PC batches, (b) score contribution of bath PD11 at the timepoint 24, and (c) score contribution of bath PD12 at the timepoint 15.

complete agreement with the conclusions of our previous work.<sup>9</sup>

According to these results, it was concluded that the process faults could be detected by developing a total sugar control chart based on an in-line monitoring model. While the HPLC method (used as the reference method here) took about 30 min to provide a set of results for a sample, the NIR technique required only a few seconds to provide an estimate. The prediction accuracy and increased number of results provided by NIR spectroscopy allowed minimal manual sampling and traditional testing. Additional benefits included reduced worker exposure to alcohol mixtures and reduced environmental pollution caused by reagents.

#### 4. Conclusion

NIR spectroscopy has been applied to the *in-situ* monitoring of a laboratory-scale batch alcohol precipitation process. Key parameters of glucose, fructose, and sucrose were modeled by PLS regression based on the *in-situ* NIR spectra. The internal and external validations of the model gave the RMSECV values of 1.038, 1.390, and 2.217 mg·mL<sup>-1</sup>, and RMSEP values of 0.711, 1.547, and 3.740 mg·mL<sup>-1</sup>, for the above three saccharides, respectively. The good utility values of the models were found to predict the tendency of saccharide parameters during the batch alcohol precipitation process. The application of *in-situ* NIR spectroscopy in this study confirms that NIR spectroscopy, combined with multivariate analysis, can be successfully applied to modeling and predicting the kinetic variables of alcohol precipitation (namely, glucose, fructose, and sucrose), despite the complexity inherent in studies of HM systems. The major advantages of the proposed method in this study are the use of no toxic reagents, *in-situ* monitoring, and high cost-effectiveness. The availability of alcohol precipitation descriptors in real-time could be applied to both laboratory and industrial scales. Future extensions of this work include the application of the same methods to laboratory and industrial samples, determination of other quality parameters, such as tannins or other specific impurities, and the combination of process parameters with NIR spectroscopy to better understand the processes. This will allow nondestructive, accurate, and near-real-time determination,

monitoring, and control of the whole batch process to be realized.

#### Conflict of Interest

The authors declare that there is no conflict of interest.

#### Acknowledgments

This work was supported by the State Administration of Traditional Chinese Medicine of Zhejiang Province Project (No. 2015ZQ022) and the Zhejiang TCM Health Science and Technology Project (No. 2015KYB110).

#### References

1. X. C. Gong, S. S. Wang, H. B. Qu, "Comparison of two separation technologies applied in the manufacture of botanical injections: Second ethanol precipitation and solvent extraction," *Ind. Eng. Chem. Res.* **50**, 7542–7548 (2011).
2. G. Eggleston, J. R. Vercellotti, "Degradation of sucrose, glucose and fructose in concentrated aqueous solutions under constant PH conditions at elevated temperature," *J. Carbohydr. Chem.* **19**, 1305–1318 (2000).
3. M. Coca, M. T. Garcia, G. Gonzalez, M. Pena, J. A. Garcia, "Study of colored components formed in sugar beet processing," *Food Chem.* **86**, 421–433 (2004).
4. B. Y. Yang, R. Montgomery, "Alkaline degradation of fructofuranosides," *Carbohydr. Res.* **280**, 47–57 (1996).
5. F. S. Asghari, H. Yoshida, "Acid-catalyzed production of 5-hydroxymethyl furfural from D-fructose in subcritical water," *Ind. Eng. Chem. Res.* **45**, 2163–2173 (2006).
6. I. Severin, C. Dumont, A. Jondeau-Cabaton, V. Graillot, M. C. Chagnon, "Genotoxic activities of the food contaminant 5-hydroxymethylfurfural using different *in vitro* bioassays," *Toxicol. Lett.* **192**, 189–194 (2010).
7. G. Schmourlo, R. R. Mendonca-Filho, C. S. Alviano, S. Cost, "Screening of antifungal agents using ethanol precipitation and bioautography of medicinal and food plant," *J. Ethnopharmacol.* **96**, 563–568 (2005).
8. J. Huang, G. Kaul, C. S. Cai, R. Chatlapalli, P. Hernandez-Abad, K. Ghosh, A. Nagi, "Quality by design case study: An integrated multivariate

- approach to drug product and process development," *Int. J. Pharm.* **382**, 23–32 (2009).
9. H. X. Huang, H. B. Qu, "In-line monitoring of alcohol precipitation by near-infrared spectroscopy in conjunction with multivariate batch modeling," *Anal. Chim. Acta.* **707**, 47–56 (2011).
  10. J. Xu, R. Q. Yue, J. Liu, H. M. Ho, T. Yi, H. B. Chen, Q. B. Han, "Structural diversity requires individual optimization of ethanol concentration in polysaccharide precipitation," *Int. J. Biol. Macromol.* **67**, 205–209 (2014).
  11. J. Židková, J. Cabálková, "Determination of saccharides in fruit juices by capillary electrophoresis and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry," *J. Mass Spectrom.* **36**, 417–421 (2001).
  12. B. Lai, M. R. Plan, M. P. Hodson, J. O. Krömer, "Simultaneous determination of sugars, carboxylates, alcohols and aldehydes from fermentations by high performance liquid chromatography," *Fermentation* **2**, 6-1–6-7 (2016).
  13. L. Zhao, A. M. Chanon, N. Chattopadhyay, I. E. Dami, J. J. Blakeslee, "Quantification of carbohydrates in grape tissues using capillary zone electrophoresis," *Front. Plant Sci.* **7**(818), 1–14 (2016).
  14. J. P. Higgins, S. M. Arrivo, G. Thurau, R. L. Green, W. Bowen, A. Lange, A. C. Templeton, D. L. Thomas, R. A. Reed, "Spectroscopic approach for on-line monitoring of particle size during the processing of pharmaceutical nanoparticles," *Anal. Chem.* **75**, 1777–1785 (2003).
  15. U.S. Department of Health and Human Services *et al.*, "Guidance for industry PAT: A framework for innovative pharmaceutical development, manufacturing, and quality assurance," <http://www.fda.gov/downloads/Drugs/Guidances/ucm070305.pdf> (2004).
  16. M. Blanco, J. Coello, H. Iturriaga, S. Maspoch, C. Pezuela, "Critical review near-infrared spectroscopy in the pharmaceutical industry," *Analyst* **123**, 135R–150R (1998).
  17. E. W. Ciurczak, J. K. Drennen, *Pharmaceutical and Medical Applications of Near-Infrared Spectroscopy*, Marcel Dekker, New York (2002).
  18. X. Xiao, J. F. Ma, F. H. Ge, X. D. Zhang, H. H. Yang, Q. L. Liang, Y. M. Wang, G. A. Luo, "Application of near-infrared spectroscopy for the rapid analysis of *Lonicerae Japonicae Flos* solution extracted by water," *J. Innov. Opt. Health Sci.* **7**(4), 1350069-1–1350069-9 (2014).
  19. C. Y. Wu, J. S. Chen, M. R. Li, Y. Wu, X. S. Liu, "Improvement of NIR models for quality parameters of leech and earthworm medicines using outlier multiple diagnoses," *J. Innov. Opt. Health Sci.* **11**(1), 1750009-1–1750009-15 (2018).
  20. P. J. Brimmer, F. A. DeThomas, J. W. Hall, Method development and implementation of near-infrared spectroscopy in industrial manufacturing process, in *Near-Infrared Technology in the Agriculture and Food Industries*, pp. 199–214, American Association of Cereal Chemists, Inc., St. Paul, MN (2001).
  21. D. A. Burns, E. W. Ciurczak, *Handbook of Near-Infrared Analysis*, 2nd Edition, Practical Spectroscopy Series, Vol. 35, Marcel Dekker, New York (2001).
  22. J. Luybaert, D. L. Massart, Y. V. Heyden, "Near-infrared spectroscopy applications in pharmaceutical analysis," *Talanta* **72**, 865–883 (2007).
  23. ASTM, "ASTM Standard E2537-08: Standard guide for the application of continuous quality verification to pharmaceutical and biopharmaceutical manufacturing," ASTM International, West Conshohocken (2008).
  24. J. Huang, S. R. Torres, M. Moshgbar, "Practical considerations in data pre-treatment for NIR and Raman spectroscopy," *Pharm. Rev.* **6**, 116–127 (2010).
  25. T. Hu, T. T. Li, L. Nie, L. X. Zang, H. C. Zang, Y. Z. Zeng, "Rapid monitoring the water extraction process of *Radix Paeoniae Alba* using near infrared spectroscopy," *J. Innov. Opt. Health Sci.* **10**(3), 1750002-1–1750002-8 (2017).
  26. H. J. Chae, S. W. Chae, D. H. Yun, K. S. Keum, S. K. Yoo, H. R. Kim, "Prevention of bone loss in ovariectomized rats: The effect of *Salvia miltiorrhiza* extracts," *Immunopharmacol. Immunotoxicol.* **26**(1), 135–144 (2004).
  27. A. H. Liu, L. Li, M. Xu, Y. H. Lin, H. Z. Guo, D. A. Guo, "Simultaneous quantification of six major phenolic acids in the roots of *Salvia miltiorrhiza* and four related traditional Chinese medicinal preparations by HPLC-DAD method," *J. Pharma. Biomed. Anal.* **41**(1), 48–56 (2006).
  28. J. Z. Xu, J. Shen, Y. Y. Cheng, H. B. Qu, "Simultaneous detection of seven phenolic acids in Danshen injection using HPLC with ultraviolet detector," *J. Zhejiang Univ., Sci. B* **9**(9), 728–733 (2008).
  29. R. Rosipal, N. Krämer, Overview and recent advances in partial least squares, *Subspace, Latent Structure and Feature Selection*, Lecture Notes in Computer Science, Vol. 3940, pp. 34–51, Springer, Berlin (2006).
  30. R. N. M. Páscoa, J. A. Lopes, J. L. F. C. Lima, "In situ near infrared monitoring of activated dairy sludge wastewater treatment processes," *J. Near Infrared Spectrosc.* **16**(4), 409–419 (2008).
  31. P. Geladi, B. Kowalski, "Partial least squares regression: A tutorial," *Anal. Chim. Acta.* **185**, 1–17 (1986).

32. J. Mantanusa, E. Ziémons, P. Lebruna, E. Rozeta, R. Klinkenberg, B. Strel, B. Evrard, Ph. Hubert, "Active content determination of non-coated pharmaceutical pellets by near infrared spectroscopy: Method development, validation and reliability evaluation," *Talanta* **80**, 1750–1757 (2010).
33. W. L. Li, H. B. Qu, "Characterization of herbal powder blends homogeneity using near-infrared spectroscopy," *J. Innov. Opt. Health Sci.* **7**(6), 1450004-1–1450004-7 (2014).
34. A. Nordon, A. Mills, R. T. Burn, F. M. Cusick, D. Littlejohn, "Comparison of non-invasive NIR and Raman spectrometries for determination of alcohol content of spirits," *Anal. Chim. Acta.* **548**, 148–158 (2005).
35. L. J. Xie, X. Q. Ye, D. H. Liu, Y. B. Ying, "Quantification of glucose, fructose and sucrose in bayberry juice by NIR and PLS," *Food Chem.* **114**, 1135–1140 (2009).