Improved spectral model building for using ATR-FTIR spectroscopy to measure solution concentration during cooling crystallization

Jingxiang Liu[a,b], Tao Liu [b,c,*], Yan Cui [b,c], Xiaojing Pei [b,c]

a School of Marine Electrical Engineering, Dalian Maritime University, Dalian, 116026, P. R. China
b Key Laboratory of Intelligent Control and Optimization for Industrial Equipment of Ministry of Education, Dalian University of Technology, Dalian 116024, P. R. China
c Institute of Advanced Control Technology, Dalian University of Technology, Dalian 116024, P. R. China
** Corresponding author, e-mail: tliu@dlut.edu.cn*

Abstract: To ensure in-situ measurement accuracy on the solution concentration during cooling crystallization process, an improved spectral model building method is proposed in this paper for application of the attenuated total reflection Fourier transform infrared (ATR-FTIR) spectroscopy. The traditional partial least-squares (PLS) model is modified into a functional regression form for expressing the relationship between the input variables of in-situ measured spectra and the output of solution concentration, so as to address the issue of insufficient samples for model calibration with respect to the high-dimension of spectral variables for measurement. The widely used DB4 wavelet functions are taken as basis functions to approximate the smooth functions in the proposed functional regression model, by virtue of their multi-scale and orthogonal properties to procure good accuracy. Accordingly, a parameter estimation algorithm named wavelet partial-least-squares is proposed for the spectral model calibration. The application to measure the solution concentration of L-glutamic acid (LGA) cooling crystallization process well demonstrates the effectiveness and merit of the proposed method.

Keywords: wavelet functional partial least-squares, in-situ measurement by ATR-FTIR spectroscopy, cooling crystallization process, spectral calibration model.

1. INTRODUCTION

Cooling crystallization has been widely used in modern industries for product separation and purification, such as fine-chemicals, pharmaceuticals, sugars and salts. To procure good crystal products with desired size, shape, and purity, it is key important to regulate the solution concentration (SC) during a cooling crystallization process. It is therefore required to measure SC in real time with good accuracy, to facilitate the process control and monitoring. Since a densitometer was proposed to measure SC in real time (Garside and Mullin, 1996), a few techniques based on the refractive index, densitometry, conductometry, or calorimetry were explored for measuring SC (Helt and Larson, 1977; Monnier, Fevotte, Hoff, 1997; Hermanto, Phua, Chow, and Tan, 2013). With the rapid development of spectroscopy technology in the past two decades, the corresponding instruments for in-situ measurement have been developed for crystallization process monitoring, including the attenuated total reflectance Fourier transform infrared (ATR-FTIR) spectroscopy, near infrared (NIR) spectroscopy, and ATR-UV/vis. Owing to the sensitivity to the main components and their contents in crystal solution, the ATR-FTIR has been increasingly applied for in-situ measurement of various cooling crystallization processes (Kadam, Mesbah, Windt and Kramer, 2011; Nagy and Braatz, 2012; Zhang, Liu, Wang, Liu and Jiang, 2017).

However, it was pointed out that there could exist notable prediction error on SC in using the ATR-FTIR spectroscopy for in-situ measurement of crystallization process (Gherras and Fevotte, 2012; Zhang, Liu, Wang, Liu and Jiang, 2017). The main reason lies with inaccurate spectral calibration model building. Due to the fact that the spectral absorbance exhibits certain nonlinearity under different SC levels, the classical linear regression modeling methods, e.g., principal component regression (PCR), partial least-squares (PLS), could not guarantee prediction accuracy in the undersatuated zone (USZ) or metastable zone (MSZ) during crystallization (Borissova, Khan, Mahmud, Roberts, Andrews, Dallin, Chen, and Morris, 2009). Besides, the dimension of spectral variables is typically high in a specified wavelength range to reflect the sensitivity of main components in the crystal solution, e.g., there are almost 1000 wavenumbers that cover the range of main absorption peaks of L-glutamic acid (LGA) solution. In contrast, only a small number of different SC levels could be prepared in engineering practice for spectral calibration model building. In other words, only a small number of samples could be collected for such model building in the presence of high- dimensional inputs variables of spectral wavenumbers. Multivariate statistical methods were explored to deal with very limited data for process modeling problem in the recent years (Camacho, Picó, and Ferrer, 2009; Liu, Liu, Chen, and Qin, 2018; Tulsyana, Garvinb and Undey, 2019). Principal component regression,

partial least-squares (PLS), multiple linear regression, and support vector machine methods were developed for spectral calibration (Barla, Kumar, Nalluri, Gandhi and Venkatesh, 2014; Zhang, Liu, Wang, Liu and Jiang, 2017; Brestrich, Rüdt, Büchler and Hubbuch, 2018; Mu, Liu, Liu, Xia and Yu, 2019). It was found that PLS is superior to the other methods in terms of the commonly used index of minimum prediction error, based on a limited number of samples for model building (Brestrich, Rüdt, Büchler and Hubbuch, 2018). However, the nonlinear modeling problem remains open as yet.

Note that the measured spectra were usually regarded as discrete data for analysis when applying the classical multivariate statistical methods, whereas the continuous characteristics of these spectral variables were neglected. In contrast to multivariate statistical methods based on finite dimensional vectors of discrete samples, functional data analysis (FDA) emphasizes the smoothness of a fitting curve, and therefore is more suitable for analyzing non-stationary time series with unequally spaced observations (Ferraty and Vieu, 2006; Mears, Nørregard, Sin, Krist, Stocks and Albaek, 2016). Good applications of FDA for modeling continuous-time process dynamics can be found in the literature (Ramsay and Silverman, 2005; Cuevas, 2014; He and Zhu, 2016; Qin and Chiang, 2019). Among the developed functional analysis methods, the classical spline functions were early studied in the references (Frank, 1990; Wold, 1992). Artificial neural network (ANN) was adopted in combination with the partial least squares (PLS) to handle the nonlinear issue (Zhu, Chen, He and Yu, 2017). A comparative study on multiple linear regression, principal component regression, PLS, polynomial PLS regression, spline PLS regression and ANN, was conducted in the reference (Roman, Ravilya and Ekaterina, 2007). Recently, wavelet analysis has received increasing attentions owing to the excellent properties of orthogonality, compact support and multi-resolution (Chen, Yang and Wei, 2012; Chen, Li and Racic, 2018). It was recognized that wavelet functions could be well used to conduct feature extraction and tackle nonlinear fitting problem (Tsakiroglou, Sygouni and Aggelopoulos, 2010; Cuevas, 2014). However, wavelet functions were not explored for building spectral calibration model to predict SC, which motivates this study.

In this paper, a novel functional based calibration model building method for engineering application of ATR-FTIR spectroscopy is proposed to improve the measurement accuracy on SC during cooling crystallization process, based on using wavelet functions in combination with PLS (abbreviated by WFPLS). The in-situ measured spectral curves are closely approximated by continuous functions based on constructing wavelet basis functions. Then the proposed WFPLS algorithm is presented to estimate the model parameters for spectral calibration with desirable accuracy. For clarity, the paper is organized as follows: The proposed functional modeling method is detailed in Section 2, including problem description, and the proposed WFPLS algorithm. In Section 3, experimental verification on measuring the solution concentration of LGA cooling crystallization process is shown to demonstrate the effectiveness and merit of the proposed method. Finally, some conclusions were drawn in Section 4.

2. PROPOSED FUNCTIONAL CALIBRATION MODEL BUILDING

2.1 Problem description

To clearly explain the salient problems involved with spectral calibration for in-situ measurement of SC, Figure 1 shows the measured spectra for 6 different SC levels of LGA solution at the same temperature of 46 °C before crystallization for illustration.

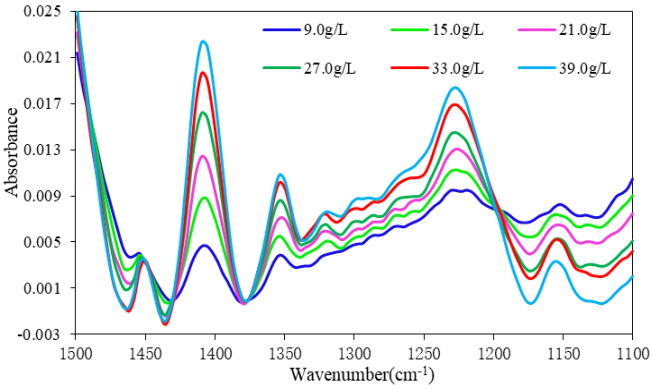


Figure 1. Illustration of measured spectra for different SC levels of LGA solution at the same temperature for crystallization

It is seen that the measured spectra cover a wavenumber range of 1100-1500 cm^{-1}, which reflects the main absorption peaks of LGA functional groups, i.e., there are at least 400 wavenumbers that should be taken as the input variables for spectral calibration model building. In contrast, no more than 40 different SC levels could be prepared for spectral calibration experiments in practice, due to the fact that the highest SC of LGA is below 40 g/L and the lowest SC is above 1 g/L that could be precisely measured by ATR-FTIR spectroscopy. Hence, it is challengeable to establish an accurate spectral calibration model with such high-dimension spectral variables in terms of very limited SC samples. Moreover, there exist obvious nonlinear correlations between spectral absorption peaks. For example, the lowest SC of 9 g/L corresponding to the blue line in Figure 1 shows the lowest absorption peak around the wavenumber of 1400 cm^{-1}, but the highest peak around the wavenumber of 1150 cm^{-1}. In contrast, the highest SC of 39 g/L corresponding to the cyan line in Figure 1 shows the adverse effect at the referred wavenumbers. This phenomenon indicates that the spectral absorbance has notable nonlinearity with respect to different SC levels. Most of the existing spectral calibration models as aforementioned, however, belong to linear modeling methods, such that the prediction accuracy could not be guaranteed under the above nonlinearity.

The main task of this work is therefore aimed at establishing a functional calibration model to overcome the above problems, such that the in-situ measurement accuracy could be guaranteed for application to cooling crystallization process. To this end, a novel functional WFPLS model is detailed in the following section.

2.2 Proposed WFPLS

Suppose that the measured spectra (input data) are denoted by $\mathbf{X}=[\boldsymbol{x}_1,\cdots,\boldsymbol{x}_N]^{\mathrm{T}}\in\Re^{N\times M}$, where $\boldsymbol{x}_i\in\Re^M$ is the i-th spectrum, N the number of spectra, and M the number of wavelengths in each spectrum. The output data of SC are denoted as $\boldsymbol{y}\in\Re^N$.

Generally, the traditional PLS modeling method could be expressed by

$$\begin{aligned}&\max\ <\mathbf{X}\boldsymbol{w},\mathbf{Y}\boldsymbol{v}>\\&s.t.\ \ \boldsymbol{w}^{\mathrm{T}}\boldsymbol{w}=1,\ \ \boldsymbol{v}^{\mathrm{T}}\boldsymbol{v}=1\end{aligned}\tag{1}$$

where $\boldsymbol{w}$ and $\boldsymbol{v}$ are the loading vectors with respect to $\mathbf{X}$ and $\mathbf{Y}$, respectively.

Considering that each of the measured NIR spectra is a functional curve, it is therefore expressed by a function denoted by $s_n(t)$, where t is the functional argument. Correspondingly, the prediction of the scalar variable y_n is predicted by the function $s_n(t)$ instead of the raw high-dimensional wavenumber variables, such that the referred high-dimensional problem and spectral nonlinearity associated with model building can be subtly circumvented.

Accordingly, a functional extension of PLS prediction on $\boldsymbol{y}_i$ is based on constructed functions $s_i(t)$ where t denotes the function argument. The loading vector $\boldsymbol{w}$ is therefore replaced by a loading function $w(t)$. Since $s_i(t)$ and $w(t)$ are continuous, the sum of the inner product becomes integral. In this way, each spectrum can be approximated by a continuous function, and the matrix $\mathbf{X}$ becomes a function vector $\boldsymbol{s}(t)=[s_1(t),\cdots,s_N(t)]$. Thus Eq.(1) can be expressed as follows.

$$\begin{aligned}&\max\ <\boldsymbol{s}(t)w(t),\mathbf{Y}\boldsymbol{v}>\\&s.t.\ \int_\Omega w(t)w(t)dt=1,\ \ \boldsymbol{v}^{\mathrm{T}}\boldsymbol{v}=1\end{aligned}\tag{2}$$

For implementation, the spectral function $s_i(t)$ is constructed as a linear combination of basis functions, i.e.,

$$s_i(t)=\sum_{k=1}^{K}c_{ik}\varphi_k(t)=\boldsymbol{c}_i^{\mathrm{T}}\boldsymbol{\varphi}(t)\tag{3}$$

where $\boldsymbol{\varphi}(t)=[\phi_1(t),\cdots,\phi_K(t)]^{\mathrm{T}}$ is the basis function vector, $\boldsymbol{c}_i=[c_{i1},\ldots,c_{iK}]^{\mathrm{T}}$ the corresponding coefficient vector, K the number of basis function, $i=1,\cdots,N$.

Hence, the function vector $\boldsymbol{s}(t)$ can be expressed by

$$\boldsymbol{s}(t)=\mathbf{C}\boldsymbol{\varphi}(t)\tag{4}$$

where $\mathbf{C}=[\boldsymbol{c}_1,\cdots,\boldsymbol{c}_N]^{\mathrm{T}}\in\Re^{N\times K}$ is the coefficient matrix.

Similarly, the loading function $w(t)$ can also be described as a linear combination of basis functions,

$$w(t)=\sum_{k=1}^{K}b_k\varphi_k(t)=\boldsymbol{b}^{\mathrm{T}}\boldsymbol{\varphi}(t)\tag{5}$$

where $\boldsymbol{b}=[b_1,\cdots,b_K]^{\mathrm{T}}$ is the coefficient vector.

With the functional expressions of Eq.(4) and Eq.(5), the following equation is easily obtained,

$$<\boldsymbol{s}(t)w(t),\mathbf{Y}\boldsymbol{v}>=(\int_\Omega\boldsymbol{s}(t)w(t)dt)^{\mathrm{T}}\mathbf{Y}\boldsymbol{v}\tag{6}$$

owing to

$$\int_\Omega\boldsymbol{s}(t)w(t)dt=\int_\Omega\mathbf{C}\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{b}dt=\mathbf{C}\mathbf{J}\boldsymbol{b}\tag{7}$$

where $\mathbf{J}=\int_\Omega\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)dt$.

When orthogonal basis functions are adopted, $\mathbf{J}$ becomes an identity matrix. The object in Eq.(6) becomes $\boldsymbol{b}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}\mathbf{Y}\boldsymbol{v}$. Similarly, the constraint $\int_\Omega w(t)w(t)dt=1$ becomes

$$\int_\Omega\boldsymbol{b}^{\mathrm{T}}\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{b}dt=\boldsymbol{b}^{\mathrm{T}}\boldsymbol{b}=1\tag{8}$$

Correspondingly, the proposed WFPLS model becomes

$$\begin{aligned}&\max\ <\mathbf{C}\boldsymbol{b},\mathbf{Y}\boldsymbol{v}>=\max\ \boldsymbol{b}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}\mathbf{Y}\boldsymbol{v}\\&s.t.\ \ \boldsymbol{b}^{\mathrm{T}}\boldsymbol{b}=1,\ \ \boldsymbol{v}^{\mathrm{T}}\boldsymbol{v}=1\end{aligned}\tag{9}$$

Note that the proposed model in Eq.(9) is an extension of the traditional PLS model, so the classical NIPALS algorithm can be used to estimate the model parameters. Accordingly, the output prediction could be obtained by

$$\hat{\mathbf{Y}}=\mathbf{C}\boldsymbol{\Theta}\tag{10}$$

$$\boldsymbol{\Theta}=\sum_{i=1}^{A}\boldsymbol{b}_i^{*}\boldsymbol{r}_i^{\mathrm{T}}\tag{11}$$

$$\boldsymbol{b}_i^{*}=\prod_{j=1}^{i-1}(\mathbf{I}-\boldsymbol{b}_j\boldsymbol{p}_j^{\mathrm{T}})\boldsymbol{b}_i\tag{12}$$

where $\hat{\mathbf{Y}}$ is the predicted output matrix, $\boldsymbol{p}_j$ the jth loading vector, $\boldsymbol{r}_i$ the ith regression vector, $\mathbf{I}$ the identity matrix with proper dimension, A the number of retained latent factors.

For the convenience of implementation, the orthogonal wavelet functions are taken as basis functions in the proposed WFPLS model, such as DB4 wavelet. To perform spectral calibration, the first step is to determine the approximation functions for the raw spectral curves. The multiscale active approximation algorithm recently developed (Liu, Chen and Wang, 2020) is used herein for simplicity, where the multi-scale and orthogonal DB4 wavelet is adopted. The wavelet basis functions are determined in terms of the approximation error as user specified in practice. Typically, a predefined threshold of fitting error is used for this purpose.

For real application, when a new spectral measurement denoted by $\boldsymbol{x}_{\mathrm{new}}$ comes, the corresponding function $s_{\mathrm{new}}(t)$ is determined by

$$s_{\mathrm{new}}(t)=\boldsymbol{c}_{\mathrm{new}}^{\mathrm{T}}\boldsymbol{\varphi}(t)\tag{13}$$

$$\boldsymbol{c}_{\mathrm{new}}=(\boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{T}})^{-1}\boldsymbol{\Phi}\boldsymbol{x}_{\mathrm{new}}\tag{14}$$

where $\boldsymbol{c}_{\mathrm{new}}$ is the approximation vector. Then the prediction of SC is obtained by

$$\hat{\boldsymbol{y}}=\boldsymbol{c}_{\mathrm{new}}^{\mathrm{T}}\boldsymbol{\Theta}\tag{15}$$

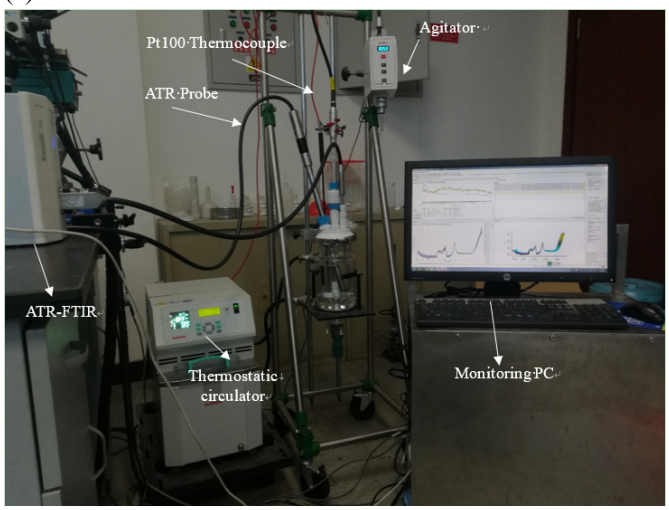
where $\hat{\boldsymbol{y}}$ is the predicted SC, $\boldsymbol{\Theta}$ is defined in Eqs. (11) and (12).

Therefore, SC can be predicted in real time based on the measured spectra data.

3. CASE STUDY ON MEASURING COOLING CRYSTALLIZATION PROCESS

The experimental set-up and its schematic of cooling crystallization is shown in Figure 2 (a) and (b).

(a)



(b)

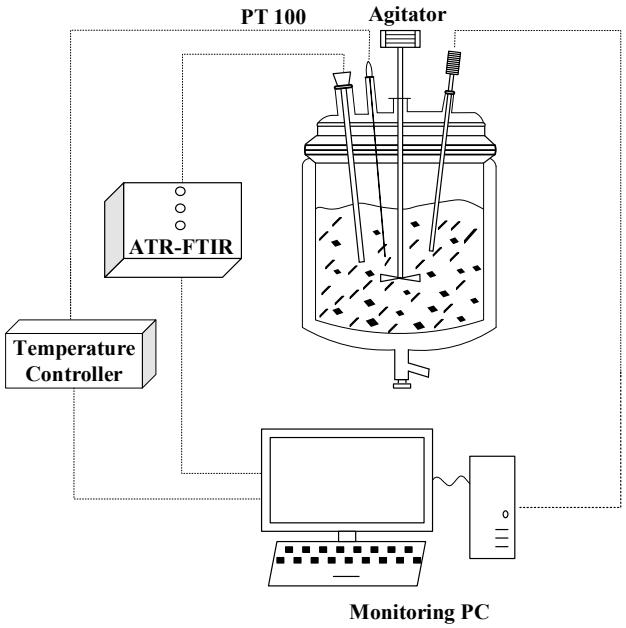


Figure 2. Experimental set-up for LGA cooling crystallization: (a) external view; (b) schematic diagram.

The crystallizer consists of a 1-liter (L) glass jacketed reactor, a thermostatic circulator (product no. Julabo-CF41), a PT100 thermometer, and a PTFE four-paddle agitator. A diamond ATR immersion probe connected via AgX Fiber as the internal reflectance element attached to the FTIR spectroscopy was used to collect the absorbance spectra of LGA solution. The ATR-FTIR spectra were monitored by a software named ReactIR15 made by Mettler-Toledo Company. The probe was set in the fine mode with the laser focused at 0 μm before experiments and set at a scanning speed of 2 m/s during experiments. All data collected by ATR-FTIR were displayed in the ReactIR 15 software packages. The raw materials for cooling crystallization are β-LGA (C5H9NO4, purity: 99%, produced by Sigma) along with distilled water as the solvent. A high-resolution analytical balance (Mettler-Toledo) with a precision of one ten thousandth was used to weigh the β-LGA samples for experiments. The stirring speed was maintained at 150 r/min for all experiments.

In the study, the experiments corresponding to six different SC levels, i.e., 9, 15, 21, 27, 33, and 39 g/L, are conducted. All spectra were measured after the LGA solute dissolved completely in each experiment. A total of 628 spectra were measured, consisting 208, 134, 114, 76, 64, and 52 spectra for each corresponding SC levels. In this study, a half of the measured data is used for training and the other half for test.

To assess the performance of constructed models, the root mean square error (RMSEP) and R^2 defined below are used,

$$RMSE = \sqrt{\frac{(\boldsymbol{y}-\hat{\boldsymbol{y}})^{\mathrm{T}}(\boldsymbol{y}-\hat{\boldsymbol{y}})}{I}} \tag{16}$$

$$R^2 = 1-\frac{(\boldsymbol{y}-\hat{\boldsymbol{y}})^{\mathrm{T}}(\boldsymbol{y}-\hat{\boldsymbol{y}})}{(\boldsymbol{y}-\overline{\boldsymbol{y}})^{\mathrm{T}}(\boldsymbol{y}-\overline{\boldsymbol{y}})} \tag{17}$$

where $\hat{\boldsymbol{y}}$ is the predicted values of SC in each experiment, and $\overline{\boldsymbol{y}}$ is the mean vector.

The proposed WFPLS model and the classical PLS model are applied for comparison. For application of the proposed method, the raw measured spectra should be converted into functions in the first place. Here, the DB4 wavelet is used and the approximation results for one representative spectrum collected from LGA solution is shown in Figure 3.

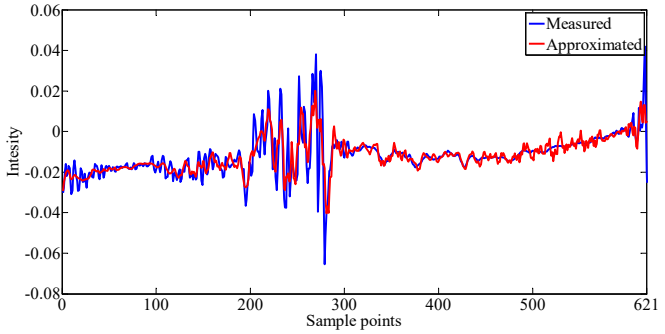


Figure 3. Approximation result for one representative spectrum collected from LGA solution by the proposed functional method.

It is seen that the high-dimension spectrum is well approximated by using 64 wavelet basis functions, especially for the front 300 points, where the measurement noise is effectively filtered out. Accordingly, the dimension of modeling data is significantly reduced. The prediction results for all the test data by the WFPLS and PLS models are shown in Figure 4 (a), where the blue line with circles denotes the measured data, black line with black crosses denotes the predicted values by PLS, and the line with red stars denotes the predicted values by WFPLS. Figure 4 (b) and (c) shows the enlarged plots corresponding to SC of 9 g/L and 33 g/L, respectively.

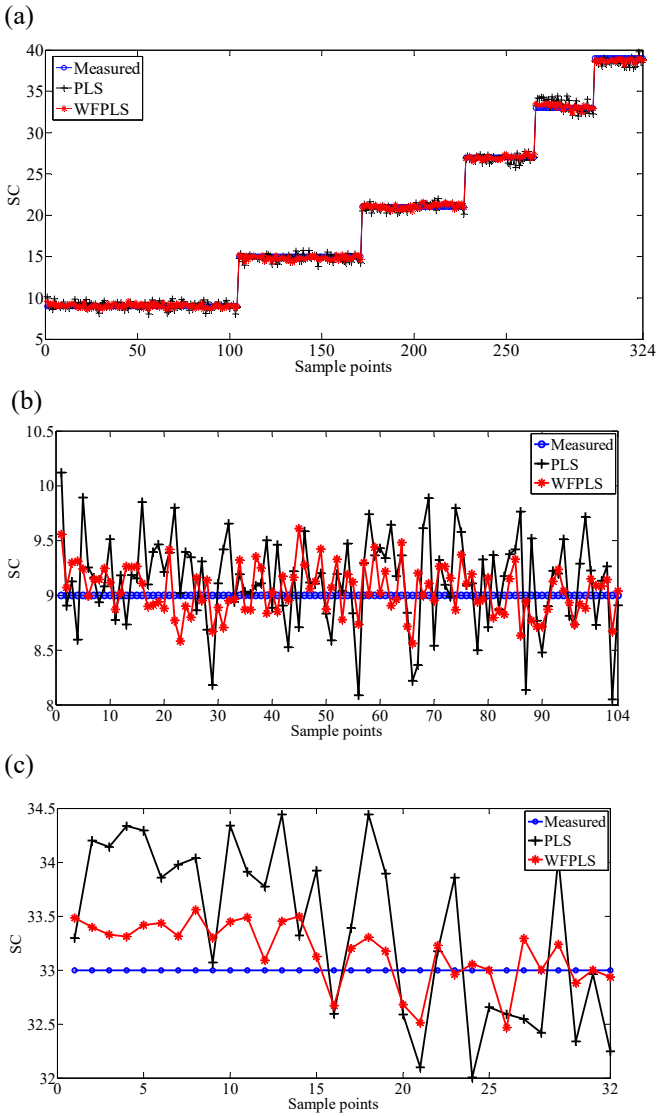


Figure 4. Prediction results for cooling crystallization process corresponding to: (a) all different SCs; (b) the enlarged view for SC=9 g/L; (c) the enlarged view for SC=33 g/L

It is seen that the prediction results obtained by the proposed WFPLS are obviously closer to the true values. The corresponding RMSE and R^2 values for these two methods are shown in Table 1. It can be seen that the R^2 values of the two proposed functional methods are over 0.999, showing obvious advantages over the traditional PLS method.

Table 1 Comparison of two methods for testing data

Methods	RMSE	R^2
MPLS	0.5142	0.9973
WFPLS	0.2651	0.9993

It can be seen that the R^2 values of the proposed functional method is over 0.999 and the prediction error is obviously smaller than that of PLS, demonstrating that significantly improved prediction accuracy is obtained in comparison with the traditional PLS method.

4. CONCLUSIONS

In this paper, a novel functional calibration method named WFPLS has been proposed for applying ATR-FTIR spectroscopy to in-situ measure SC during cooling crystallization process. By modifying the traditional PLS model into a functional regression form, the spectra data are easily transformed into smooth functions, such that only a small number of concentration samples is needed for model calibration with respect to the high-dimension of spectral variables for measurement. The widely used DB4 wavelet functions are adopted as basis functions to approximate the smooth functions in the proposed functional regression model, which could procure good accuracy for model fitting and prediction against nonlinear spectral properties. The experimental measurement results on LGA cooling crystallization process well demonstrates the advantage of the proposed spectral calibration method.

ACKNOWLEDGEMENT

This work is supported in part by NSF China Grants 62003071 and 62173058, Open Funding Project from Key Laboratory of Intelligent Control and Optimization of Industrial Equipment, Ministry of Education (LICO2021TB01), the Talent Project of Revitalizing Liaoning (XLYC1902030), and the Fundamental Research Funds for the Central Universities of China (DUT21LAB113).

REFERENCES

Barla, V.S., Kumar, R., Nalluri, V.R., Gandhi, R.R., & Venkatesh, K. (2014). A practical evaluation of qualitative and quantitative chemometric models for real-time monitoring of moisture content in a fluidised bed dryer using near infrared technology. J. Near Infrared Spectrosc. 22, 221-228.

Borissova, A., Khan, S., Mahmud, T., Roberts, K.J., Andrews, J., Dallin, P., Chen, Z.P., Morris, J., (2009). In situ measurement of solution concentration during the batch cooling crystallization of L-glutamic acid using ATR-FTIR spectroscopy coupled with chemometrics, Crystal Growth & Design, 9, 692-706.

Brestrich, N., Rüdt, M., Büchler, D., Hubbuch, J., (2018). Selective protein quantification for preparative chromatography using variable pathlength UV/Vis spectroscopy and partial least squares regression. Chem. Eng. ci. 176 (2), 157-164.

Camacho, J., Picó, J., & Ferrer, A (2009). The best approaches in the online monitoring of batch processes based on PCA: does the modeling structure matter? Anal. Chim. Acta, 642, 59-69.

Chen, J.H., Yang, Y.C., Wei, T.Y., (2012). Application of wavelet analysis and decision tree in UTDR data for diagnosis of membrane filtration. Chemometr. Intell. Laboratory Systems 116, 102-111.

Chen, J., Li, G., Racic, V., (2018). A data-driven wavelet-based approach for generating jumping loads. Mech. Syst. Sig. Process. 106, 49-61.

Cuevas, A., (2014). A partial overview of the theory of statistics with functional data. J. Statist. Planning Inference 147, 1–23.

Ferraty, F., Vieu, P., (2006). Nonparametric functional data analysis: Theory and Practice. Springer, New York.

Frank, I.E., (1990). A nonlinear PLS model. Chemometrics and Intelligent Laboratory Systems, 8(2), 109-119.

Garside, J. & Mullin, J (1966). Continuous measurement of solution concentration in a crystalliser, Chemistry & Industry, 48, 2007.

Gherras, N., Fevotte, G., (2012). Comparison between approaches for the experimental determination of metastable zone width: A case study of the batch cooling crystallization of ammonium oxalate in water, Journal of Crystal Growth, 342, 88-98.

He, Y.L., Zhu, Q.X., (2016). A novel robust regression model based on functional link least square (FLLS) and its application to modeling complex chemical processes. Chem. Eng. Sci. 153 (22), 117-128.

Helt, J.E., Larson, M.A., (1977). Effects of temperature on the crystallization of potassium nitrate by direct measurement of supersaturation. AIChE J., 23 (6), 822-830.

Hermanto, M.W., Phua, A., Chow, P.S., & Tan, R.B (2013). Improved C-control of crystallization with reduced calibration effort via conductometry, Chem. Eng. Sci. 97, 126-138.

Kadam, S.S., Mesbah, A., Windt, E., Kramer, H.J.M. (2011). Rapid online calibration for ATR-FTIR spectroscopy during batch crystallization of ammonium sulphate in a semi-industrial scale crystallizer, Chemical Engineering Research and Design, 89, 995-1005.

Liu, J.X., Chen, J.H., Wang, D., (2020). Wavelet functional principal component analysis for batch process monitoring. Chemometrics and Intelligent Laboratory Systems, 196, 103897.

Liu, J.X., Liu, T., Chen, J.H., Qin, P., (2018), Novel common and special features extraction for monitoring multi-grade processes, Journal of Process Control, 66, 98-107

Liu, J.X., Liu, T., Chen, J.H., Yue, H., Zhang, F.K. & Sun, F.R. (2020) Data-driven modeling of product crystal size distribution and optimal input design for batch cooling crystallization processes. Journal of Process Control, 96, 1-14

Mears, L., Nørregard, R., Sin, G., Krist, V., Stocks, G.S.M., Albaek Kris Villez, M.O., (2016). Functional unfold principal component regression methodology for analysis of industrial batch process data. AIChE J. 62, 1986-1994.

Monnier, O., Fevotte, G., Hoff, C., (1997). Model identification of batch cooling crystallizations through calorimetry and image analysis. Chem. Eng. Sci., 52 (7), 1125-1139.

Mu, G.Q., Liu, T., Liu, J.X., Xia, L.Z., Yu, C.Y., (2019) Calibration model building for on-line monitoring of the granule moisture content during fluidized bed drying by NIR spectroscopy. Industrial & Engineering Chemistry Research, 58(16), 6476-6485

Nagy, Z.K. & Braatz, R.D (2012). Advances and new directions in crystallization control, Chem. Biomol. Eng. 3, 55-75.

Qin, S.J., & Chiang, L.H. (2019). Advances and opportunities in machine learning for process data analytics. Computers & Chemical Engineering, 161, 465-473.

Ramsay, J.O., Silverman, B.W., (2005). Functional data analysis, Springer Series in Statistics. Springer, New York.

Roman, M.B., Ravilya Z.S., Ekaterina I.L., (2007). Comparison of linear and nonlinear calibration models based on near infrared (NIR) spectroscopy data for gasoline properties prediction. Chemometrics and Intelligent Laboratory Systems 88,183-188.

Tsakiroglou, C.D., Sygouni, V., Aggelopoulos, C.A., (2010). Using multi-level wavelets to correlate the two-phase flow characteristics of porous media with heterogeneity. Chem. Eng. Sci. 65 (24), 6452–6460.

Tulsyana, A., Garvinb, C., Undey, C., (2019). Industrial batch process monitoring with limited data. Journal of Process Control, 77, 114-133.

Wang, X.Z., Roberts, K.J. & Ma, C. (2008) Crystal growth measurement using 2D and 3D imaging and the perspectives for shape control, Chem. Eng. Sci., 63, 1173-1184.

Wold, S., (1992). Nonlinear partial least squares modelling. II. Spline inner relation. Chemometrics and Intelligent Laboratory Systems,14(1-3), 71-84.

Zhang, D.J., Liu, L.D., Xu, S.J., Du, S.C., Dong, W.B., & Gong, J.B (2018). Optimization of cooling strategy and seeding by FBRM analysis of batch crystallization, J.Cryst. Growth, 486, 1-9.

Zhang, F.K., Liu, T., Wang, X.Z., Liu, J.X., Jiang, X.B., (2017). Comparative study on ATRFTIR calibration models for monitoring solution concentration in cooling crystallization. J. Cryst. Growth 459, 50-55.

Zhu, B., Chen, Z.S., He, Y.L., Yu, L.A., (2017). A novel nonlinear functional expansion based PLS (FEPLS) and its soft sensor application. Chemometrics and Intelligent Laboratory Systems,161, 108-117.