Functional Soft Sensor Based on Spectra Data for Predicting Multiple Quality Variables

JINGXIANG LIU1, DAN WANG1, (Senior Member, IEEE), JUNGHUI CHEN2, AND JIE HOU3

1School of Marine Electrical Engineering, Dalian Maritime University, Dalian 112016, China
2Department of Chemical Engineering, Chung-Yuan Christian University, Taoyuan 32023, Taiwan
3College of Automation, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Corresponding author: Junghui Chen (jason@wavenet.cycu.edu.tw)

This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant 3132019101, in part by the National Natural Science Foundation of China under Grant 51579023, Grant 61673081, and Grant 61903057, in part by the Ministry of Science and Technology, Taiwan, under Grant MOST 106-2221-E-033-060-MY3, and in part by the Chongqing Natural Science Foundation under Grant cstc2019jcyj-msxmX0129.

ABSTRACT In many complex chemical processes, such as ethanol fermentation process, prediction of the multivariate quality variables based on spectra data presents a great challenge because the dimensions of the spectra far exceed their sampling number. To address this problem, the dimension reduction of the predictors is necessary. It can be conducted either by regressing on components or by smoothing methods with basis functions. Based on the functional data analysis methods, this work introduces a novel wavelet functional partial least squares (WFPLS), which combines both of the foregoing dimension-reduction approaches. The high-dimensional spectra can be well fitted by fewer wavelet basis functions in the proposed method. Using the proposed WFPLS method does not require the measured data to be linear and to be sampled on a regular basis. It will be proved that the proposed WFPLS method can be finally transferred into the traditional PLS method in computation. By comparison with the existing prediction methods, the advantages of the proposed method are well demonstrated via a numerical case and an ethanol fermentation experiment.

INDEX TERMS Wavelet functional partial least squares, soft sensor, functional data analysis, spectra data analysis, active learning.

I. INTRODUCTION

In practice, spectrum technologies, such as middle infrared and near-infrared (NIR), have increasing applications in many chemical plants because of a significant virtues of non-destructive and no sample preparation [1]–[4]. By comparison with the traditional sensors, spectrum technologies have the merits of no invasion and thus have little effects on the production process. Meanwhile, different from the traditional measured samples that contain discrete variables, the spectra have the nature of a distribution [5], [6]. The dimension of each spectrum is very high and much larger than the number of measured samples, which is a typical undersampling problem. Thus, it is meaningful and necessary to establish a more efficient and accurate soft sensor model with respect to spectra data.

In the recent years, data-driven modeling methods achieve great applications in many field, because they have few requirement on the prior knowledge, easy operation, great generalization performance [7]–[10]. To get better final product quality and improve the industrial benefits, it is meaningful and necessary to establish an accurate online prediction model for practical applications. In the past publications, the spectral models established by multiple linear regression, principal component regression, partial least-squares (PLS), and support vector regression (SVR) were studied and compared [11]–[14]. It was found that PLS is better than the other methods as the former has the minimum prediction error. With the prior process knowledge, some process variables were found to have a great effect on the final product quality, which were used for model building by combination with NIR spectroscopy [5], [15]. However, most existing calibration modeling methods are linear models. They may not perform well in the practical case with serious nonlinear characteristics. To handle the implied nonlinear problem, kernel based methods, such as kernel PLS and SVR, is proposed to improve the prediction performance [16], [17]. However, how to choose the proper kernel functions and tune the kernel
parameters is still an open issue to be solved, hindering the practical applications of the kernel based methods.

Analytically, the measured spectra data have several characteristics as follows: 1) high dimension. There may be hundreds or thousands of variables in each measured spectra samples; 2) the limited number of data. The number of modeling samples is not adequate for establishing an accurate calibration model by compared to the variables; 3) nonlinear property. As most measured spectral data have a serious nonlinear relationship. Linear methods often fail in many practical cases. Based on the above analysis, the dimension of spectral data should be reduced before modeling. Meanwhile, the nonlinear relationship should be taken into consideration in order to obtain more accurate calibration performance [4], [15]. In most published calibration methods, the modeling data are analyzed discretely and discretely by the traditional multivariate statistical methods, without considering the implied continuous property within the variables. In the measured spectra data, the spectrum trajectories often appear as smooth curves. Therefore, functional data analysis (FDA) methods can be applied to spectra data analysis [18]–[21]. Explanatorily, FDA deals with the analysis and the theory of data that are in the form of functions [19], [20]. By comparison, FDA methods are suitable for the irregular non-stationary time series analysis and emphasize the smoothness of the variable trajectories [22], [23]. Functional regression and functional PLS models also have found applications in chemical processes [1], [24], [25], where B-spline functions were mostly applied as basis functions for computation. However, the above mentioned basis functions belong to global function, which may cause redundant computation in many situations. Assume that a curve is well approximated except only a partial local range. If another global basis function is added, the well fitted ranges will be degraded. So, more global basis functions must be added to achieve the required accuracy. To handle the above problem, the multi-scale wavelet functions are increasing used for approximation and show great advantages [26]–[28]. Moreover, a novel active strategy was applied to calculate the wavelet basis functions recently. More concise model and better modeling performance can be obtained.

In this paper, a novel wavelet functional partial least squares (WFPLS) is proposed to build soft sensors using the measured spectroscopy. Wavelet functions are applied to fit the trajectories of each sampled spectrum. To get more concise model and better performance, an active strategy is used to calculate the number and locations of the basis functions. By setting a proper threshold, the spectral curves can be well approximated by the proposed method no matter how heavy the nonlinear curve is. Then the functional partial least squares model is performed on the determined basis functions. In this way, the parameters of the prediction model are obviously reduced, and the nonlinear problem can be easily solved. The main contents of this paper include the following sections. Firstly, the preliminary knowledge about the traditional PLS and the active strategy are introduced in Section 2. Then the proposed WFPLS method, as well as the online application and some discussions, are given in Section 3. For illustration, a numerical case and an experimental study on prediction the multiple quality variables of the ethanol fermentation process by using spectra are discussed in Section 4. Finally, Some conclusions are drawn in the end.

II. PRELIMINARIES

To enhance understanding, the traditional PLS and the active strategy to calculate the basis functions are firstly reviewed in the following.

A. THE TRADITIONAL PLS

Suppose the measured spectra are denoted as $X = [x_1, \ldots, x_N]^T \in \mathbb{R}^{N \times M}$, where $x_i \in \mathbb{R}^M$ is the $i$th spectrum, $N$ and $M$ are the numbers of spectra and wavelengths in each spectrum, respectively. The quality data are denoted as $Y \in \mathbb{R}^{N \times J}$, where $J$ is the number of quality variables. The traditional PLS method performs regression on the collinear variables by projecting to a lower dimensional latent space [29]. The PLS model can be built as follows.

$$\max_{s.t. \ w^T w = 1, \ v^T v = 1} \ < Xw, Yv >$$

(1)

where $w$ and $v$ are the loading vectors with respect to $X$ and $Y$, respectively. The above optimization problem can be solved by the Lagrange multiplier technique as follows

$$\begin{cases} X^T Y Y^T Xw = \lambda_1 w \\ Y^T X X^T Yv = \lambda_2 v \end{cases}$$

(2)

where $\lambda_1$ and $\lambda_2$ are the eigen values, $w$ and $v$ are the corresponding eigen vectors, respectively. Then the latent vectors $t$ and $u$ can be calculated as

$$\begin{cases} t = Xw = \lambda_1 w \\ u = Yv \end{cases}$$

(3)

The modelling data can be updated as $X - tp^T$ and $Y - btq^T$, where $p = X^T t / (t^T t)$ is the loading vector, $b = u^T t / (t^T t)$ is the regression coefficient between $u$ and $t$. In this way, the components can be determined one by one iteratively. More details about the traditional PLS model can be found in [29]–[31].

B. ACTIVE STRATEGY TO DETERMINE WAVELET BASIS FUNCTIONS

In this work, DB4 wavelet is applied for illustration and application owing to its good properties of multi-scale, orthogonal and compact support. To get more delicate model and better modelling performance, an active learning strategy is recently proposed to determine the numbers and locations of wavelet basis functions [4], [20].

The key idea of the active strategy is to improve the fitting performance in the location corresponding to the maximum error in each step. For clear, an illustration figure is
shown in Figure 1. The top curve $e^2$ represents the mean approximation error vector calculated by the selected basis functions.

$$e_m^2 = \sum_{n=1}^{N} (x_{n,m} - \hat{x}_{n,m})^2 / N$$

where $\hat{x}_{n,m}$ is the fitting value of $x_{n,m}$ by the previously selected basis functions, the error vector $e^2 = [e^2_1, \ldots, e^2_M]^T$.

The location with respect to the maximum error is recognized and marked by the red point. Correspondingly, a wavelet basis function is firstly added and the error values around are updated as zeros. Similarly, the location with respect to the maximum error of the updated curve $\tilde{e}^2$ can be sequentially determined and marked by brown point. The second wavelet basis function can be added correspondingly and the error curve is sequentially updated. Repeat the above steps until the approximation error meet the predefined threshold. More details about the active strategy can be found in [4], [20], which is thus omitted here.

**III. THE PROPOSED SOFT SENSOR METHOD**

In this section, the proposed WFPLS is detailed and discussed. Sequentially, the online applications of the proposed WFPLS methods is detailed, and some comments are given for better understanding.

**A. WFPLS METHOD**

The functional extension of PLS prediction of quality $y_i$ is based on the functions $s_i(t)$ with $t$ the function argument. To extend the idea of PLS, the loading vector $w$ is replaced by a loading function $w(t)$. Since the function $s_i(t)$ and $w(t)$ are continuous, integral operation is thus involved. In this way, each spectrum can be approximated by a continuous function, and the matrix $X$ becomes a function vector $s(t) = [s_1(t), \ldots, s_N(t)]$. Then Eq.(1) can be expressed as follows.

$$\max < s(t)w(t), Yv >$$

$$\text{s.t.} \int_{\Omega} w(t)w(t)dt = 1, \quad v^T v = 1 \quad (5)$$

Note that the other symbols are the same as defined before. For computation, the spectral function $s_i(t)$ can be described as a linear combination of the wavelet basis functions

$$s_i(t) = \sum_{k=1}^{K} c_{ik} \varphi_k(t) = c_i^T \varphi(t) \quad (6)$$

where $\varphi(t) = [\phi_1(t), \ldots, \phi_K(t)]^T$ and $c_i = [c_{i1}, \ldots, c_{iK}]^T$ are the basis function and the coefficient vectors, respectively, and $K$ is the number of basis function, $i = 1, \ldots, N$. The coefficient $c_i$ can be computed by the least-square (LS) method as follows.

$$\min \sum_{j=1}^{M} (x_{ij} - \sum_{k=1}^{K} c_{ik} \varphi_k(t_j))^2 \quad (7)$$

Or in the matrix form

$$\min x_i - \Phi c_i^T (x_i - \Phi^T c_i) \quad (8)$$

where $\Phi = [\varphi_i(s_j)] \in \Re^{N \times K}$. Then the criterion is minimized by the solution,

$$c_i = (\Phi \Phi^T)^{-1} \Phi x_i \quad (9)$$

Thus, the function vector $s(t)$ can be expressed by

$$s(t) = \Phi \varphi(t) \quad (10)$$

where $\Phi = [c_1, \ldots, c_N]^T \in \Re^{N \times K}$ is the coefficient matrix.

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

$$w(t) = \sum_{k=1}^{K} b_k \varphi(t) = b^T \varphi(t) \quad (11)$$

where $b = [b_1, \ldots, b_K]^T$ is the coefficient vector. Considering the functional expressions of Eq.(10) and Eq.(11), the following equation is easily obtained

$$< s(t)w(t), Yv > = \left( \int_{\Omega} s(t)w(t)dt \right)^T Yv \quad (12)$$

Owing to

$$\int_{\Omega} s(t)w(t)dt = \int_{\Omega} \Phi \varphi(t) \varphi^T(t)b dt = \Phi C b \quad (13)$$

where $J = \int_{\Omega} \varphi(t) \varphi^T(t)dt$. The object in Eq.(12) becomes $b^T \Phi^T \Phi C^T Yv$. Similarly, the constraint $\int_{\Omega} w(t)w(t)dt = 1$ becomes

$$\int_{\Omega} b^T \Phi \varphi(t) \varphi^T(t)b dt = b^T J b = 1 \quad (14)$$
In this way, the WFPLS model becomes

$$\max < CJb, Yv > = \max b^T J^T C^T Y v$$
$$s.t. \ b^T J b = 1, \ v^T v = 1 \quad (15)$$

If orthogonal basis functions are used, $J$ will be an identity matrix with the dimension of $K \times K$. WFPLS in Eq. (15) can thus be equivalent to the traditional PLS, given by

$$\max < Cb, Yv > = \max b^T C^T Y v$$
$$s.t. b^T b = 1, \ v^T v = 1 \quad (16)$$

Kindly note that the expression in Eq. (16) is exactly the same as the traditional PLS by replacing the raw data $X$ as the coefficients $C$. This conclusion is tenable only when the orthogonal basis functions involved. Otherwise, we cannot obtain the formula in Eq. (16). In this way, the traditional NIPALS algorithm can be applied to calculating the model parameters. Then the quality variables can be predicted by

$$\hat{Y} = C\Theta \quad (17)$$
$$\Theta = \sum_{i=1}^{A} b_i^* r_i^T \quad (18)$$
$$b_i^* = \prod_{j=1}^{i-1} (1 - b_j p_j^T) b_i \quad (19)$$

where $\hat{Y}$ is the predicted quality matrix, $p_j$ is the $j$th loading vector, $r_i$ is the $i$th regression vector, $I$ is the identity matrix with the proper dimension, $A$ is the number of retained latent factors.

### B. ONLINE APPLICATION

For the online application, when a new vector $x_{new}$ is available, it should be approximated by the wavelet basis functions firstly. Based on the trained basis functions, the corresponding function $s_{new}(t)$ is determined by

$$s_{new}(t) = c_{new}^T \varphi(t) \quad (20)$$
$$c_{new} = (\Phi \Phi^T)^{-1} \Phi x_{new} \quad (21)$$

where $c_{new}$ is the approximation vector. Then the prediction of the quality variable can be calculated by

$$\hat{y} = c_{new}^T \Theta \quad (22)$$

where $\hat{y}$ is the predicted quality vector, $\Theta$ is defined in Eqs. (18) and (19). Thus, the quality variables can be predicted in real-time based on the easily measured NIR spectra data.

For better understanding, a brief flow chart of the off line training and online application of the proposed WFPLS is given in Figure 2.

### C. COMMENTS AND DISCUSSION

As stated in the proposed method, the high-dimensional spectral data are approximated by wavelet basis functions and modeled in the functional space directly. The proposed WFPLS has the following implied advantages:

1. The high-dimensional data can be approximated by dozens of basis functions. In some sense, the information and properties of the raw data are inherited by the approximation coefficients. Therefore, the approximation step in the proposed method can be taken as a feature extraction or dimension reduction.

2. Using the multi-scale wavelet functions as basis functions, the orthogonal property is inherited in the chosen basis functions.

3. The proposed method has no extra requirement for the process data. No matter how serious nonlinear and non-Gaussian the spectrum is, the proposed method can still obtain a satisfactory approximation result.

The high-dimensional and nonlinear problems can be well handled by the proposed wavelet functional method. Then the quality can be accurately predicted based on the approximation coefficients. Without fewer assumptions on the modeling data, the proposed functional method is easy to use in practice.

### IV. ILLUSTRATION EXAMPLES

#### A. NUMERICAL CASE

The numerical data with continuous inputs are generated as follows

$$s_i(t) = \alpha_1 P_1(t \mid \mu_1, \sigma_1) + \alpha_2 P_2(t \mid \mu_2, \sigma_2) + \alpha_3 P_3(t \mid \mu_3, \sigma_3) \quad (23)$$

where $t$ is the variable uniformly sampled in the range of $[0, 5]$, $P_i(t \mid \mu_i, \sigma_i)$, $i = 1, 2, 3$, follow the Gaussian distributions, where $\mu_i$ and $\sigma_i$ are the mean and the variance, and

$$\begin{align*}
\alpha_1 &= 0.35, & \mu_1 &= 0.5, & \sigma_1 &= 0.5 \\
\alpha_2 &= 0.35, & \mu_2 &= 2, & \sigma_2 &= 0.5 \\
\alpha_3 &= 0.4, & \mu_3 &= 3, & \sigma_3 &= 0.2
\end{align*} \quad (24)$$

![Flow chart about the off line training and online application of the proposed WFPLS method.](image2.png)
By defining a mapping function that reflects nonlinear output response as follows

$$
\begin{align*}
    h_1(t) &= (-t^3 + 4t^2 + 7t)/5 \\
    h_2(t) &= (-0.5t^3 + 2.5t^2 + 5t)/5 \\
    h_3(t) &= 2h_1(t) + 1 \\
    h_4(t) &= 2h_2(t) - 2 \\
    h_5(t) &= 2h_3(t) + h_4(t)
\end{align*}
$$

(25)

The output data is generated by $y_{ij} = \int_{\Omega} s_i(t)h_j(t)dt$, where $\Omega = [0, 5]$, $i = 1, 2, \ldots, 80$, $j = 1, 2, \ldots, 5$. The overall simulated data are $X \in \mathbb{R}^{80 \times 500}$ and $Y \in \mathbb{R}^{80 \times 5}$, where 40 samples for model training and the other 40 samples for testing. Note that the Gaussian noise $N(0, 0.04)$ is involved in the numerical case.

In this work, the root mean square error for the prediction (RMSEP) and $R^2$ are used to demonstrate the prediction performance,

$$
RMSEP = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2
$$

(26)

$$
R^2 = 1 - \frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2}
$$

(27)

where $y_i$, $\hat{y}_i$, and $\bar{y}$ denote the $i$th quality value, its prediction and its mean value, respectively.

Using the proposed WFPLS method, a total of 25 wavelet functions are determined for computation. So that the dimension of modeling data in WFPLS method can be obviously reduced compared to the raw data $X \in \mathbb{R}^{80 \times 500}$ used in the traditional methods. Figure 3 demonstrates the approximation results of the numerical data, where the blue solid line is the simulated curve, and the red dot dash line is approximated by the proposed WFPLS method. The results well demonstrate the approximation effectiveness against the noise.

The proposed WFPLS, SVR and PLS are used to predict the quality variables for comparison in this case. For the application of SVR, the Gaussian kernel is used to handle the nonlinear problem of the simulated data. Here the parameters of each method are tuned by the traditional cross-validation methods. Not list all, the prediction results for the third quality variable $Y_3$ by the proposed WFPLS, SVR, and PLS are shown in Figure 4 (a), (b), and (c), respectively. The blue lines with circles are the true simulated values, and the red lines with stars are the values predicted by different methods. To make clear comparisons, the predictions for a local region are zoomed in a subplot in the original plot.

As shown in Figure 4 (a), the predicted values by WFPLS exactly overlap the true values. The values predicted by the SVR and PLS methods have some deviations to the true values more or less (Figure 4 (b) and (c)). For better comparison, the RMSEP and $R^2$ values by the three methods are shown in Table 1 and 2. The prediction results indicate the proposed WFPLS is much better than SVR and PLS, with the minimum RMSEP values and the $R^2$ values roughly equal to 1.
In the saccharomyces cerevisiae glycolysis process, the ethanol and carbon dioxide are produced by the reaction, with a small amount of energy released. In this process, the concentration of glucose, the concentration of ethanol and the biomass are very important parameters. This work is based on a real-time detection system of NIR technology as shown in Figure 5 (a), and its schematic diagram is given in Figure 5 (b). In the experiment, the capacity of the fermenter is 2.5 L, NaOH is used to control the pH values in the fermenter, which is measured by the pH meter, the reactor temperature is measured by the thermocouple PT100, the temperature is controlled to the predefined value by a heating device and cold water, and the electric stirring paddle is used to agitate the fermentation evenly. During the fermentation process, the NIR spectrum analyzer is used to measure the acquisition of the fermented liquid under different wavelength absorbance values in real-time. Then the measured spectra data are sent to the monitoring computer. The prediction model for the glucose concentration, ethanol concentration, and biomass can then be established based on the measured data.

Saccharomyces cerevisiae is used as the strain in this experiment. The seeds need to be cultured before the fermentation experiment is conducted. The Fourier NIR analyzer (TALYS-ASP531) produced by ABB company and its matching immersion diffuse reflection probe are used to measure the NIR spectra data. For modeling, the wavelength range is $4,790–12,000 \text{ cm}^{-1}$, the instrument resolution is $16 \text{ cm}^{-1}$, the number of spectral scans is 64, and the detector gain is 237.84. The samples which are centrifuged and diluted are measured every 0.5 h. Then liquid chromatography (water) is used to measure the glucose concentration, gas chromatography (Agilent 6890 Series GC System) is used to measure the ethanol concentration, and the microplate reader (Multiskan Ascent) is used to measure the biomass (OD) in the fermentation process. The measured spectral data are shown in Figure 6.

A total of 6 batches of the ethanol fermentation process is conducted. Among them, 5 batches with 138 samples are used for training, and the other one batch with 26 samples is used for testing. Using the proposed function method, the raw high-dimensional data can be approximated by only dozens of wavelet basis functions. It is statically enough to train the model by 138 samples. The proposed WFPLS, SVR and PLS are used to predict the quality variables, including the glucose concentration, the ethanol concentration, and biomass, in the ethanol fermentation process. Similarly, the Gaussian kernel is used to handle the nonlinear problem within the
spectra data. The parameters of each method are tuned by the traditional cross-validation methods. The prediction results using different methods for the ethanol concentration in the experiment are given in Figure 7. The blue lines with circles are the true measured values, and the red lines with stars are the predicted values. It can be found that the values predicted by the proposed WFPLS method are closer to the true values than those predicted by SVR and PLS methods.

RMSEP and $R^2$ by the three methods are shown in Table 3 and 4. It can be seen that the prediction results for the glucose concentration and the ethanol concentration obtained by the proposed WFPLS are much better than those obtained by the SVR and PLS, with the minimum RMSEP values and the maximum $R^2$ values.

**TABLE 3. RMSEP for the ethanol fermentation process.**

| Methods | Glucose | Ethanol | OD  |
|---------|---------|---------|-----|
| WFPLS   | 2.0790  | 1.0924  | 0.1926 |
| SVR     | 3.2428  | 1.5833  | 0.1220 |
| PLS     | 3.3763  | 2.1585  | 0.1483 |

**TABLE 4. $R^2$ for the ethanol fermentation process.**

| Methods | Glucose | Ethanol | OD  |
|---------|---------|---------|-----|
| WFPLS   | 0.9898  | 0.9898  | 0.9897 |
| SVR     | 0.9741  | 0.9699  | 0.9931 |
| PLS     | 0.9719  | 0.9440  | 0.9899 |

**V. CONCLUSION**

A novel WFPLS method is proposed for quality prediction based on high-dimensional spectral data in this manuscript. In the proposed method, wavelet basis functions are used to fit the spectra data. Thus, the high-dimensional spectral data are modeled as smooth curves and analyzed directly in the functional space. By proper deviations, the WFPLS modeling problem can be converted into the traditional PLS one because of the orthogonal property of wavelet functions. The online application and discussion are detailed in this work. The numerical case is used to demonstrate the effectiveness and advantages of the proposed WFPLS by comparison with the traditional PLS and SVR methods. The experimental ethanol fermentation process is used to testify the practical feasibility and effectiveness of the proposed method. For future study, the applications of WFPLS for various process conditions will be further researched.

**REFERENCES**

[1] K. C. Sales, F. Rosa, P. N. Sampaio, L. P. Fonseca, M. B. Lopes, and C. R. C. Calado, “In situ near-infrared (NIR) versus high-throughput mid-infrared (MIR) spectroscopy to monitor biopharmaceutical production,” *Appl. Spectrosc.*, vol. 69, no. 6, pp. 760–772, Jun. 2015.

[2] G. L. Yang, W. H. Lu, Y. Lin, J. Z. Luo, C. B. Wang, R. Meder, P. Warburton, and R. J. Arnold, “Monitoring water potential and relative water content in Eucalyptus camaldulensis using near infrared spectroscopy,” *J. Tropical Forest Sci.*, vol. 29, pp. 121–128, Jan. 2017.

[3] R. Moscetti, F. Raponi, S. Ferri, A. Colantonii, D. Monarca, and R. Massantini, “Real-time monitoring of organic apple (var. Gala) during hot-air drying using near-infrared spectroscopy,” *J. Food Eng.*, vol. 222, pp. 139–150, Apr. 2018.

[4] J. X. Liu, T. Liu, G. Q. Mu, and J. H. Chen, “Wavelet-based calibration model building of NIR spectroscopy for in-situ measurement of granule moisture content during fluidized bed drying,” *Chem. Eng. Sci.*, vol. 223, Nov. 2020, Art. no. 115867, doi: 10.1016/j.ces.2020.115867.

[5] R. Kona, H. Qu, R. Mattes, B. Jancsik, R. M. Fahmy, and S. W. Hoag, “Application of in-line near infrared spectroscopy and multivariate batch modeling for process monitoring in fluid bed granulation,” *Int. J. Pharmaceutics*, vol. 452, nos. 1–2, pp. 63–72, Aug. 2013.

[6] V. S. Barla, R. Kumar, V. R. Nalluri, R. R. Gandhi, and K. Venkatesh, “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 Infr. Spectrosc.*, vol. 22, no. 3, pp. 221–228, Jun. 2014.

[7] J. Zhu, Y. Wang, D. Zhou, and F. Gao, “Batch process modeling and monitoring with local outlier factor,” *IEEE Trans. Control Syst. Technol.*, vol. 27, no. 4, pp. 1552–1565, Jul. 2019.

[8] S. Zhang, C. Zhao, and F. Gao, “Incipient fault detection for multiphase batch processes with limited batches,” *IEEE Trans. Control Syst. Technol.*, vol. 27, no. 1, pp. 103–117, Jan. 2019.

[9] Y. Liu, C. Yang, Z. Gao, and Y. Yao, “Ensemble deep kernel learning with application to quality prediction in industrial polymerization processes,” *Chemometrics Intell. Lab. Syst.*, vol. 174, pp. 15–21, Mar. 2018.

[10] H. Yu, C. Guo, Z. Shen, and Z. Yan, “Output feedback spatial trajectory tracking control of underactuated unmanned undersea vehicles,” *IEEE Access*, vol. 8, pp. 42924–42936, 2020.

[11] L. Song, W. Huang, X. Han, and J. Mazundar, “Real-time composition monitoring using support vector regression of laser-induced plasma for laser additive manufacturing,” *IEEE Trans. Ind. Electron.*, vol. 64, no. 1, pp. 633–642, Jan. 2017.

[12] Y. Liu, K. Liu, J. Yang, and Y. Yao, “Spacial-neighborhood manifold learning for nondestructive testing of defects in polymer composites,” *IEEE Trans. Ind. Informat.*, vol. 16, no. 7, pp. 4639–4649, Jul. 2020.

[13] X. Yuan, Y. Wang, C. Yang, Z. Ge, Z. Song, and W. Gui, “Weighted linear dynamic system for feature representation and soft sensor application in nonlinear dynamic industrial processes,” *IEEE Trans. Ind. Electron.*, vol. 65, no. 2, pp. 1508–1517, Feb. 2018.

[14] J. Wu, H. Cheng, Y. Liu, B. Liu, and D. Huang, “Modeling of adaptive multi-output soft-sensors with applications in wastewater treatments,” *IEEE Access*, vol. 7, pp. 161887–161898, 2019.

[15] G. Mu, T. Liu, J. Liu, L. Xia, and C. Yu, “Calibration model building for online monitoring of the granule moisture content during fluidized bed drying by NIR spectroscopy,” *Ind. Eng. Chem. Res.*, vol. 58, no. 16, pp. 6476–6485, Apr. 2019.
[16] X. Huang and L. Xia, “Improved kernel PLS combined with wavelength variable importance for near infrared spectral analysis,” Chemometric Intell. Lab. Syst., vol. 168, pp. 107–113, Sep. 2017.

[17] X. Huang, D.-S. Cao, Q.-S. Xu, L. Shen, J.-H. Huang, and Y.-Z. Liang, “A novel tree kernel support vector machine classifier for modeling the relationship between bioactivity and molecular descriptors,” Chemometric Intell. Lab. Syst., vol. 120, pp. 71–76, Jan. 2013.

[18] J. O. Ramsay, “When the data are functions,” Psychometrika, vol. 47, no. 4, pp. 379–396, Dec. 1982.

[19] J. O. Ramsay and B. W. Silverman, Functional Data Analysis (Springer Series in Statistics), 2nd ed. New York, NY, USA: Springer, 2005.

[20] J. Liu, J. Chen, and D. Wang, “Wavelet functional principal component analysis for batch process monitoring,” Chemometric Intell. Lab. Syst., vol. 196, Jan. 2020, Art. no. 103897.

[21] J. Chen and J. Liu, “Derivation of function space analysis based PCA control charts for batch process monitoring,” Chem. Eng. Sci., vol. 56, no. 10, pp. 3289–3304, May 2001.

[22] F. Ferraty and P. Vieu, Nonparametric Functional Data Analysis: Theory and Practice. New York, NY, USA: Springer, 2006.

[23] L. Mears, R. Nørregård, G. Sin, K. V. Gernaey, S. M. Stocks, M. O. Albæk, and K. Villez, “Functional unfold principal component regression methodology for analysis of industrial batch process data,” AIChE J., vol. 62, no. 6, pp. 1986–1994, Jun. 2016.

[24] P. T. Reiss and R. T. Ogden, “Functional principal component regression and functional partial least squares,” J. Amer. Stat. Assoc., vol. 102, pp. 984–996, Sep. 2007.

[25] A. M. Aguilera, M. C. Aguiler-Morillo, and C. Preda, “ Penalized versions of functional PLS regression,” Chemometric Intell. Lab. Syst., vol. 154, pp. 80–92, May 2016.

[26] B. R. Bakshi, “Multiscale PCA with application to multivariate statistical process monitoring,” AIChE J., vol. 44, no. 7, pp. 1596–1610, Jul. 1998.

[27] B. R. Bakshi and G. Stephanopoulos, “Wave-net: A multiresolution, hierarchical neural network with localized learning,” AIChE J., vol. 39, no. 1, pp. 57–81, Jan. 1993.

[28] F. M. Dwight and J. O. Kraig, Elements of Wavelets for Engineers and Scientists. Hoboken, NJ, USA: Wiley, 2003.

[29] S. Wold, A. Ruhe, H. Wold, and W. J. Dunn, III, “The collinearity problem in linear Regression. The partial least squares (PLS) approach to generalized inverses,” SIAM J. Sci. Stat. Comput., vol. 5, no. 3, pp. 735–743, Sep. 1984.

[30] P. Geladi and B. R. Kowalski, “Partial least-squares regression: A tutorial,” Anal. Chim. Acta, vol. 185, pp. 1–17, Jan. 1986.

[31] A. Högstedsson, “PLS regression methods,” J. Chemometrics, vol. 2, no. 3, pp. 211–228, Jun. 1988.