

# **Wavelength selection-based nonlinear calibration for transcutaneous blood glucose sensing using Raman spectroscopy**

Narahara Chari Dingari  
Ishan Barman  
Jeon Woong Kang  
Chae-Ryon Kong  
Ramachandra R. Dasari  
Michael S. Feld

# Wavelength selection-based nonlinear calibration for transcutaneous blood glucose sensing using Raman spectroscopy

Narahara Chari Dingari,\* Ishan Barman,\* Jeon Woong Kang, Chae-Ryon Kong, Ramachandra R. Dasari, and Michael S. Feld†

Massachusetts Institute of Technology, G. R. Harrison Spectroscopy Laboratory, Laser Biomedical Research Center, Cambridge, Massachusetts 02139

**Abstract.** While Raman spectroscopy provides a powerful tool for noninvasive and real time diagnostics of biological samples, its translation to the clinical setting has been impeded by the lack of robustness of spectroscopic calibration models and the size and cumbersome nature of conventional laboratory Raman systems. Linear multivariate calibration models employing full spectrum analysis are often misled by spurious correlations, such as system drift and covariations among constituents. In addition, such calibration schemes are prone to overfitting, especially in the presence of external interferences that may create nonlinearities in the spectra-concentration relationship. To address both of these issues we incorporate residue error plot-based wavelength selection and nonlinear support vector regression (SVR). Wavelength selection is used to eliminate uninformative regions of the spectrum, while SVR is used to model the curved effects such as those created by tissue turbidity and temperature fluctuations. Using glucose detection in tissue phantoms as a representative example, we show that even a substantial reduction in the number of wavelengths analyzed using SVR lead to calibration models of equivalent prediction accuracy as linear full spectrum analysis. Further, with clinical datasets obtained from human subject studies, we also demonstrate the prospective applicability of the selected wavelength subsets without sacrificing prediction accuracy, which has extensive implications for calibration maintenance and transfer. Additionally, such wavelength selection could substantially reduce the collection time of serial Raman acquisition systems. Given the reduced footprint of serial Raman systems in relation to conventional dispersive Raman spectrometers, we anticipate that the incorporation of wavelength selection in such hardware designs will enhance the possibility of miniaturized clinical systems for disease diagnosis in the near future. © 2011 Society of Photo-Optical Instrumentation Engineers (SPIE). [DOI: 10.1117/1.3611006]

Keywords: blood glucose monitoring; Raman spectroscopy; wavelength selection; chemometrics; support vector machines.

Paper 11082R received Feb. 24, 2011; revised manuscript received May 19, 2011; accepted for publication Jun. 22, 2011; published online Aug. 5, 2011.

## 1 Introduction

Noninvasive diagnostic tools are of great importance in biomedicine as they dramatically reduce the cost and inconvenience associated with blood withdrawals and tissue biopsies. By virtue of its high chemical specificity and capability for reagentless detection of the sample constituents, Raman spectroscopy has received considerable interest in the biomedical community for detection of a cancerous lesion,<sup>1,2</sup> atherosclerotic plaque,<sup>3</sup> and diabetes monitoring.<sup>4,5</sup> However, the multitude of Raman-active components coupled with the endogenous tissue fluorescence usually makes the quantitative determination of the analyte of interest difficult. Multivariate calibration provides a powerful tool for spectroscopy-based chemical quantification by analyzing multiple measurements of the sample responses. The primary role of multivariate calibration methods is to develop a regression model connecting the measured spectral signals to specific sample properties (such as constituent concentrations)

for predicting the same properties of prospective samples. Typically, multivariate calibration methods such as ordinary least squares (OLS) and partial least squares (PLS)<sup>6,7</sup> employ the full spectral information to differentiate the analyte of interest from the spectral interferences. To further refine the predictive ability of the multivariate calibration techniques, appropriate selection of the spectral data points (wavelengths) has been investigated using various optimization tools such as simulated annealing<sup>8</sup> and genetic algorithms.<sup>9</sup> In fact, it has been demonstrated that optimal predictions are obtained by selecting only the analyte-specific spectral features in order to eliminate uninformative and spurious regions from calibration.<sup>10,11</sup> In this regard, Raman spectroscopy is particularly suitable for wavelength selection purposes due to its inherently narrow vibrational features.

However, the application of existing wavelength selection approaches has hitherto been largely restricted to linear calibration techniques. Given the (latent) nonlinear information that is often present in the spectral data, the linear additivity assumption of the basis spectra<sup>12</sup> may not be valid necessitating the introduction of nonlinear modeling. This is particularly important for biomedical applications where the sample optical properties and measurement conditions may substantially vary causing the

\*Contributed equally to this work.

†Deceased.

Address all correspondence to: Ishan Barman; MIT, G. R. Harrison Spectroscopy Laboratory, 77 Massachusetts Avenue, 6-205, Cambridge, Massachusetts 02139; Tel: 617-258-8694; Fax: 617-253-4513; E-mail: ishan@mit.edu.

inclusion of irrelevant sources of variation in linear calibration models. Clearly, such models (full spectrum or otherwise) are unable to perform accurate predictions in prospective samples.

To achieve the important goal of noninvasive blood glucose detection, our laboratory has previously studied the causality of the blood glucose levels to Raman spectral information<sup>13</sup> and addressed specific technical challenges including tissue turbidity,<sup>14</sup> autofluorescence and photobleaching,<sup>15</sup> and physiological lag between blood and interstitial fluid glucose.<sup>16</sup> This paper investigates the applicability of wavelength selection for linear and nonlinear multivariate calibration methods to further enhance the robustness and quality of the calibration models. In this work, robust implies the ability to make accurate predictions irrespective of the specific identity of the sample or subject. Using tissue phantom<sup>17</sup> and human subject datasets,<sup>18</sup> we show that even with a substantial reduction in the number of wavelengths sampled, we can arrive at calibration models of nearly equivalent prediction accuracy. Specifically, use of wavelength selection in conjunction with nonlinear support vector regression (SVR) provides improved accuracy, in comparison to PLS full spectrum analysis. Our analysis also indicates the presence of an intrinsic subset of spectral points dictated by the vibrational responses of the specific analyte and the interferents in the samples, and the multivariate calibration method used. This leads to a minimum allowable size of the spectral subset required to establish a clinically accurate model. Furthermore, we demonstrate the prospective transferability of a selected spectral subset across different human subjects while maintaining equivalent levels of prediction accuracy. This is a significant result from a calibration maintenance and transfer perspective that potentially opens avenues for robust prospective application of spectroscopic algorithms. The approach employed in this article is also sufficiently broad and general to address diagnosis of other diseases (e.g., cancer and atherosclerosis) as well as routine pharmaceutical and forensic analysis. Finally, such selection of limited wavelength subsets may provide new impetus to the development of tunable detection filter-based serial Raman acquisition systems.<sup>19,20</sup> These systems greatly alleviate the large spatial footprint drawback of standard Raman systems but suffer from longer acquisition time requirements. By employing only a fraction of the total set of wavelengths, we envision a significant reduction in the acquisition time because of the serial sampling nature of such systems. While the tunable filter-based serial Raman systems require further experimental validation from a SNR standpoint, the acquisition time reduction coupled with the intrinsic spatial footprint advantage may make them a desirable alternative in the near future.

## 2 Theoretical background

### 2.1 Wavelength Selection

Over the years, wavelength selection techniques have been demonstrated (theoretically<sup>21,22</sup> and experimentally<sup>23</sup>) to improve calibration model accuracy and robustness. As noted in the literature, wavelength (variable) selection can be viewed as a subset of the more generic process of dimension reduction.<sup>24</sup> Potentially, wavelength selection can enhance the stability of the model to the collinearity in the acquired Raman spectra as well as increase the interpretability of the relationship between the model and the sample compositions by reducing the number

of loading vectors to the chemical rank of the system.<sup>25</sup> The important consideration is how many and which of the spectral bands facilitate accurate prediction. Clearly, such selection is also dependent on the rest of the constituents in the investigated sample(s), because of the spectral overlap that may be observed in certain fingerprint regions.

For orientation, a brief summary of the existing wavelength selection methods is provided here. Although the wavelength selection methods are sufficiently broad to work in conjunction with any calibration scheme, most of the work reported in the literature is on OLS and PLS analysis. The wavelength selection procedures primarily differ from one another in the objective criterion used for measuring the optimality of selected subsets or in the search algorithm for the determination of these subsets. Algorithms, such as simulated annealing (SA) and genetic algorithms (GA), have been proposed as global optimizers capable of determining the best set of parameters and selecting well-defined spectral regions instead of single data points scattered across the spectral range. These stochastic search methods accept transitory reductions of predictive quality during an optimization procedure enabling them to escape local extrema without supervision. However, this stochastic nature (involving predictable processes as well as random actions) is also a major disadvantage in establishing a universal spectral subset since it is almost impossible to recreate identical GA or SA models.<sup>26</sup> Other methods that have been employed for feature selection include iterative variable selection,<sup>27</sup> iterative predictor weighting,<sup>28</sup> and uninformative variable elimination.<sup>29</sup> In particular, the important work of Centner et al.<sup>29</sup> proposes the addition of artificial (noise) variables prior to the development of a closed form PLS or principal component regression (PCR) model for the dataset containing both the experimental and artificial variables. Subsequently, the experimental variables that do not exhibit more importance in predictive analysis, in comparison to the artificial variables, are eliminated. The interested reader is also referred to the work of Leardi and coworkers (for example, sequential application of backward interval PLS and GA for feature selection<sup>30</sup>) for a detailed description.

Alternately, a spectral interval selection called moving window partial least squares regression (MWPLSR) has been proposed for enhancing predictive quality of calibration models.<sup>31–33</sup> MWPLSR and its variants (changeable size moving window partial least squares and searching combination moving window partial least squares), are advantageous in searching for informative spectral regions for multicomponent spectral analysis. This method applies PLS calibration models in every window that moves over the spectrum and selects informative regions on the basis of lowest sum of residual errors. Such a moving window approach based on minimization of the residue error has previously shown promising results in improving prospective prediction of analytes in mixture solutions.<sup>34</sup>

In this paper, we follow a similar scheme of selecting lower residue error wavelength regions for both linear (PLS) and nonlinear (SVR) multivariate regression models. To implement this, we construct a spectral window of size  $w$ , which starts at  $i$ th spectral channel and ends at  $(I + w - 1)$ th spectral channel. The window is progressively moved in the full spectrum range. PLS and SVR models are built and root mean squared error of validation (RMSEV) is calculated for each spectral window position by performing prediction on the validation dataset (as detailed

in Sec. 3.2). Subsequently, we plot RMSEV as a function of the spectral window position. The spectral window positions with large errors imply that the responses at these spectral channels are highly contaminated by the factors that cannot be accurately modeled using the calibration samples. It should be noted that a residue error plot selection method searches for spectral intervals (bands), as opposed to individual scattered points, based on the continuity of the spectral response in vibrational spectra [where the characteristic full width at half maximum (FWHM) is typically at least 4 to 16  $\text{cm}^{-1}$ ].<sup>31</sup> It is worth emphasizing that residue error plot-based wavelength selection was separately applied for PLS and SVR schemes. Evidently, this resulted in distinct residue error plots for the two cases leading to separate optimal wavelength subsets.

## 2.2 Support Vector Regression

As mentioned in Sec. 1, wavelength selection schemes have been predominantly employed with linear regression models. However, the underlying assumption of linearity in the relationship between the spectra and the property of interest (analyte concentration) may not be valid under all circumstances, especially for clinical measurements. For example, in transcutaneous monitoring of blood glucose, the linearity assumption may fail due to fluctuations in process and system variables, such as changes in temperature, sampling volume and physiological glucose dynamics. Although weak nonlinearities can be modeled by the conventional linear methods such as PLS and PCR by retaining more factors than are necessitated by the chemical rank of the system, this risks the inclusion of irrelevant sources of variance and noise in the calibration model. Such incorporation of nonanalyte specific variance renders the model incapable of prospective prediction. To model the potential curved effects and to avoid overfitting of the data, support vector machines have been introduced for nonlinear classification<sup>35–37</sup> and regression (SVR).<sup>38</sup> Recently, SVR has been successfully employed for near-infrared (NIR) absorption-based concentration prediction in mixture solutions, where the acquired spectra are nonlinearly affected by temperature fluctuations.<sup>39</sup>

In SVR, the regression is performed by minimizing a cost function, which regularizes the regression coefficients and penalizes the net regression error. While reduction of large regression coefficients improves the generalization ability of the method,<sup>40</sup> minimization of the regression error (the root mean squared error of validation) ensures the development of an accurate calibration model. In SVR, this is solved with Lagrange multipliers as a constrained optimization problem and yields the following regression function:<sup>38</sup>

$$y = \sum_i (\alpha_i - \alpha_i^*) \langle x_i, x \rangle + b, \quad (1)$$

where  $x$  represents the spectral data,  $y$  is the concentration of analyte of interest,  $i$  is the index of the calibration data,  $\alpha$  and  $\alpha_i^*$  are Lagrange multipliers, and  $\langle \dots \rangle$  denotes the inner product. From Eq. (1), it is clear that each calibration data point has its own Lagrange multiplier, which decides the impact of the point on the final solution. Specifically, calibration data points that are positioned farther from the regression line (exhibiting relatively high regression errors) greatly affect the location of this line. Thus, the corresponding Lagrange multipliers are also relatively high (i.e., proportional to their regression error).

Equation (1) can be readily extended to handle nonlinear regression by substituting the inner product of the calibration and prediction spectra with a kernel function that satisfies Mercer's conditions<sup>41</sup>

$$y = \sum_i (\alpha_i - \alpha_i^*) K(x_i, x) + b. \quad (2)$$

The most widely used kernel for nonlinear regression is the radial basis function (RBF) and can be expressed as

$$K(x_i, x) = \exp\left[-\frac{\|x - x_i\|^2}{\sigma^2}\right], \quad (3)$$

where  $\sigma^2$  is the RBF kernel parameter. In addition to the kernel parameter, solving a support vector regression involves optimizing the regularization parameter, which determines the trade-off between minimizing the regression error and the regression coefficients. Further details of the analysis are described in Sec. 3.2.

## 3 Materials and Methods

### 3.1 Experimental

We have employed two data sets to investigate: (a) the relative prediction performance of wavelength selected linear and nonlinear calibration models and (b) the transferability of the wavelength selected subset from one sample to another. For (a), we employ a physical tissue model (tissue phantom) study, which focuses on glucose detection in a multicomponent mixture under controlled laboratory settings. To accomplish objective (b), a clinical dataset acquired from human subjects undergoing oral glucose tolerance tests (OGTT) is analyzed. These two data sets were originally reported in our previous publications.<sup>17,18</sup> We briefly describe the experimental methods in the following paragraphs.

For the tissue phantom data set, spectra were collected from 50 tissue phantoms containing randomly varying concentrations of two Raman active analytes, glucose, and creatinine, between 5 to 30 mM.<sup>17</sup> These samples also contained randomized concentrations of India ink and intralipid to mimic the turbidity values, absorption (0.09 to 0.18  $\text{cm}^{-1}$ ) and scattering (48.4 to 95.1  $\text{cm}^{-1}$ ), observed in human skin tissue. Spectroscopic measurements were performed on aliquots of these tissue phantoms in a fused silica cuvette by exciting with an 830 nm external cavity diode laser (Process Instruments). The back-scattered light was passed through a modified f/1.4 spectrograph (Kaiser Optical Systems, Inc.) before spectral acquisition using a liquid-nitrogen cooled CCD (Princeton Instruments).<sup>17</sup> For our analysis, we have performed curvature correction,<sup>42</sup> vertical binning, and cosmic ray removal.

For the human subject data set, transcutaneous blood glucose measurements were performed on 13 healthy Caucasian and Asian volunteers in our laboratory.<sup>18</sup> Following standard OGTT protocol, the volunteers were given 220 mL of a beverage containing 75 g of glucose before the study period. The experimental setup was similar to the one mentioned above. The laser was focused onto the forearm of the human volunteers with an average power of 300 mW and a spot size of  $\sim 1 \text{ mm}^2$ . Raman spectra were taken approximately every 5 min for each volunteer over the 2 to 3 h study period. Concomitant blood glucose

**Table 1** Summary statistics of human subject dataset.

Volunteer	Total samples	Calibration samples	Average concentration (mg/dL)	Minimum concentration (mg/dL)	Maximum concentration (mg/dL)
1	25	15	144.1	83	188
2	26	16	146.3	78	204
3	26	16	145.4	84	191
4	30	20	173.5	95	223
5	20	10	134.4	82	169
6	32	22	167.9	71	201
7	25	15	135.2	80	190
8	26	16	153.1	79	208
9	28	18	160.1	70	209
10	25	15	110.7	69	142
11	29	19	121.9	85	167
12	31	21	139.2	68	198
13	27	17	159.2	69	201

measurements were performed every 10 min from extracted blood samples, and spline interpolation was used to correlate the measured blood glucose concentrations with the spectra collected at the intermediate time points. Our human subject study was approved by the Massachusetts Institute of Technology Committee on the Use of Humans as Experimental Subjects. Prior to their inclusion in the OGTT study, informed consent was obtained from all subjects. We note that the power levels used in the study, while on the higher side, did not cause any discomfort during the test or exhibit any skin damage afterward, except in one volunteer who developed a small blister.

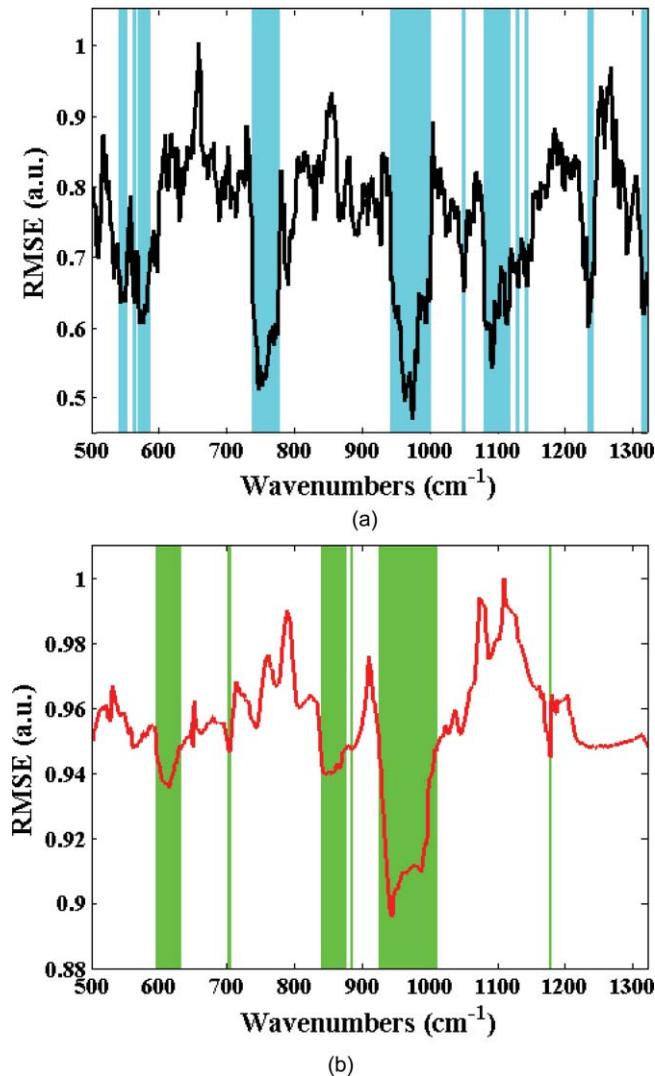
For our analysis below, datasets from volunteers exhibiting impaired glucose tolerance profiles (due to the risk of spurious correlations with quenched fluorescence levels) have been excluded. In addition, a student's *t*-test employing Mahalanobis distance function was used to reject spectra with 95% probability of being spectral outliers ( $p < 0.05$ ).<sup>43</sup> A summary of relevant statistics for the human subject study is provided in Table 1. Thirteen human subject data sets are considered for our analysis. Table 1 displays the number of calibration data points, and the average, minimum, and maximum glucose concentrations (mg/dL) for each subject.

### 3.2 Data Analysis

In the tissue phantom study, 20 prediction samples are randomly chosen from the entire data set and kept aside for prospective application. The creation of an independent prediction set is a standard approach used to mitigate and/or test for the presence of spurious correlations. Subsequently, the remaining 30 tissue

phantoms are randomly split into 20 calibration samples and 10 validation samples. The moving window approach is used to generate a regression vector from the calibration sample spectra corresponding to the specific window position. This regression vector is then used on the validation samples to obtain a RMSEV value. The window is subsequently moved, as mentioned in Sec. 2.1, over the full spectrum to construct the residue error plot as a function of the moving window position. Figure 1 shows the residue error plots calculated from a representative partition of calibration and validation sets using PLS and SVR schemes, respectively. From the residue plots, spectral points with lowest computed RMSEV are selected for developing the final regression vectors. For our analysis, we selected 100 to 900 spectral points (in increments of 100), where one spectral point roughly corresponds to  $1.45 \text{ cm}^{-1}$ . The final regression vector is generated from the 30 tissue phantoms constituted by the calibration and validation data sets for each subset of spectral points. This regression vector is prospectively used on the corresponding prediction set and root-mean-square error of prediction (RMSEP) is calculated for the specific subset of spectral points. It is worth mentioning that the prediction set comprises the spectral information only at the points selected to create the regression vector. To ensure the reproducibility of the prediction results, 100 iterations are performed to obtain an average RMSEP.

For our investigation, different window sizes from 10 to 20 spectral points are selected. This size range corresponds to the FWHM of the prominent bands in the glucose Raman spectrum. Assigning window sizes beyond this range exhibited little or adverse effect on the resultant residue lines, as also noted by



**Fig. 1** Residue error plots (solid curves) calculated for (a) PLS and (b) SVR calibration models in the tissue phantom dataset using a 20 spectral point window. The set of 300 spectral points that exhibit the minimum residue errors are highlighted in cyan and green for PLS and SVR, respectively.

Jiang et al.<sup>31</sup> Wavelength selection using the above protocol is performed for both PLS and SVR calibration. PLS models are created based on the number of loading vectors that provide the least error in cross-validation using in-house code based on the seminal algorithms of Haaland and Thomas.<sup>7</sup> A standard recommendation for PLS calibration is to incorporate at least three times the number of samples as the rank of the calibration model. To satisfy this criterion, 20 samples are chosen for calibration (as six loading vectors provided the least error in cross validation). The SVR calculations are carried out using the widely used LIBSVM toolbox originally developed by Chang and Lin<sup>44</sup> (accessible at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>). Prior to wavelength selection for SVR processing, the Raman spectra are linearly scaled by dividing each spectrum by the maximum intensity value to prevent skewed effect arising from the larger pixel intensity values. Here, we have used the RBF kernel function [Eq. (3)] to enable nonlinear scaling of the acquired dataset. The optimal model parameters  $C$  and  $\sigma^2$  are obtained by em-

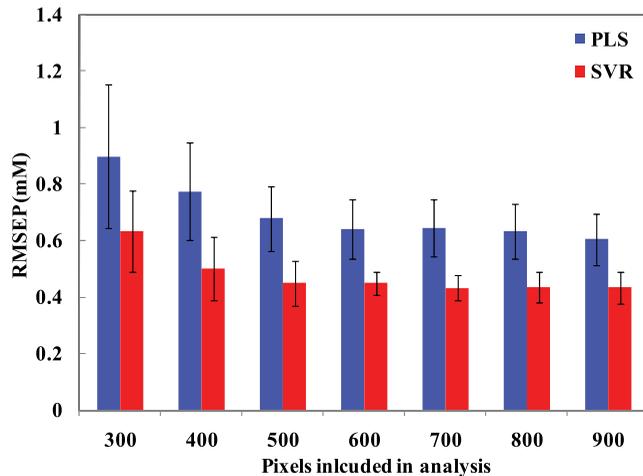
ploying a grid search algorithm in the range of 1 to 10000 ( $C$ ) and 0.01 to 10 ( $\sigma^2$ ), respectively. In addition, numerical binning of adjacent pixel intensities is performed to explore the possibility of further reduction of sampled wavelengths in conjunction with the wavelength selection approach.

Similar analysis steps are also followed for the human subject dataset. From a representative human volunteer data set (volunteer A) all but five data points are used to develop the wavelength-selected regression model. The developed regression model is then applied on the remaining five data points of volunteer A to evaluate the RMSEV as a function of the moving window position and subsequently for the construction of the residue error plots. To obtain enhanced robustness in the wavelength subset selection, we perform 100 iterations by repartitioning the calibration and validation data sets of volunteer A. The residue error plots from these iterations are added to form a cumulative error plot as a function of wavelength. By computing the cumulative validation errors using different sized spectral subsets, we find that the 300 spectral point subset provides the optimal trade-off between the number of sampled spectral points and prediction error. In other words, further reduction of number of spectral points appears to significantly compromise the predictive capability of the model. The set of 300 spectral points with the least cumulative error are then selected for prospective application in the other human volunteer datasets. For clarity, we henceforth represent the remaining volunteers as volunteer  $B_i$ , where  $i$  is the index of the volunteer. For volunteer  $B_i$ , the data points are split into calibration and prediction. The regression models are constructed on the calibration data points of volunteer  $B_i$  using only the 300 spectral points selected from volunteer A. These models are subsequently used to estimate the glucose concentrations of the prediction data points in volunteer  $B_i$ . The resultant prediction errors provide a true measure of the prospective applicability of the selected spectral subset, i.e., the transferability of the selected points across human subjects. It is worth emphasizing that we are assessing the transferability of the selected spectral subset across human subjects not that of the calibration model (regression vector) itself. The loading vectors for PLS analysis are optimized for each volunteer based on leave-one-out cross-validation on the calibration data points. Similarly, SVR optimization is also carried out for  $C$  and  $\sigma^2$  in the range of 1 to 10000 and 0.01 to 10, respectively. It is worth mentioning that for both the human volunteer as well as the tissue phantom study, the spectra were directly used for development of the multivariate calibration models without any data manipulation such as removal of background fluorescence.

## 4 Results and Discussion

### 4.1 Tissue Phantom Study

Wavelength selection was performed on the tissue phantom dataset using both PLS and SVR models. Figure 2 shows the results of prospective prediction for glucose obtained with PLS (blue) and SVR (red) models, where the lengths of the bars are proportional to the average RMSEP and the associated error bars represent the standard deviation of the RMSEP over 100 iterations. Figure 2 provides a comparative estimate of predictive performance of calibration models corresponding to the selection of minimum residue wavelength subsets of different sizes,



**Fig. 2** Bar plot showing comparative performance of wavelength-selected PLS (blue) and SVR (red) calibration models in tissue phantoms for glucose prediction. The lengths of the bars are proportional to the average RMSEP and the associated error bars represent the standard deviation of the RMSEP over 100 iterations.

ranging from 300 to 900 spectral points. It is evident that the SVR calibration models outperform the PLS calibration models with identical size of wavelength subset in regard to prospective prediction. For example, given a spectral subset size of 300 points, the mean prediction errors for glucose are observed to be 0.89 and 0.63 mM for the PLS and SVR models, respectively. In fact, the SVR model with 300 spectral points provides equivalent levels of prediction accuracy as the full spectrum PLS model (mean RMSEP of 0.6 mM) ( $p$ -value = 0.54 indicating the absence of statistically significant differences). We also observe that, as the spectral subset size is initially decreased from 900 to 500 spectral points, the change in prediction errors are not statistically significant ( $p$ -value > 0.05 for both cases). However, with further decrease of the size of the spectral subset from 500 to 300 spectral points, the prediction errors show a perceptible rise ( $p$ -value <  $10^{-4}$  for both PLS and SVR). On further reduction to 200 and 100 spectral points, the prediction error exhibits a substantially steeper rise to 1.13 and 1.64 mM for PLS and 1.73 and 2.6 mM for SVR, respectively (not shown in Fig. 2 due to their excessively high magnitude). Based on these results, one can infer how many spectral points provide relevant information specific to the analyte of interest and the spectral interferences in the sample(s). Particularly for our tissue phantom study, it would appear that the informative regions necessary for accurate glucose prediction constitute about a third of the full spectrum.

In conjunction with RMSEP determination, we have also evaluated the relative predictive determinant (RPD) metric to classify the overall prediction quality of the individual calibration models for the two Raman-active analytes (glucose and creatinine). Briefly, RPD is defined as the ratio of the standard deviation of the reference concentration in the sample population ( $\sigma_R$ ) to the standard error of prediction (the standard deviation of the differences between predicted and reference values) ( $\sigma_{R-P}$ )

$$RPD = \frac{\sigma_R}{\sigma_{R-P}}. \quad (4)$$

**Table 2** Evaluation of RPD.

Spectral points	RPD (glucose)		RPD (creatinine)	
	PLS	SVR	PLS	SVR
100	3.9	1.6	5.9	2.4
200	6.8	5.1	10.3	4.5
300	8.5	12.2	11.4	8.2
400	9.3	13.5	14.5	11.5
500	10.2	14.9	15.1	14.5
600	11.0	15.8	16.1	17.2
700	11.6	16.1	16.7	20.4
800	11.9	16.5	16.8	22.4
900	12.4	16.9	17.1	24.2

Typically, a RPD value of five is considered to be good for quality control while a value larger than 6.5 is acceptable for process monitoring. A calibration model, with a RPD value higher than eight, may be used for any application. From Table 2, we observe that both PLS and SVR models including at least 300 spectral points show excellent prediction quality. On further reduction of the number of spectral points included to 200, we find that glucose RPD values obtained using PLS and SVR are 6.8 and 5.1. When only 100 points are considered, these values show a substantive fall to 3.9 (PLS) and 1.6 (SVR), respectively. Based on these results, we can infer that a minimum of 300 spectral points is probably necessary for building a reasonable model for glucose predictions, although the specific scheme (PLS or SVR) plays a substantive role in the determination of size of the optimal wavelength subset. To this end, RPD provides a useful tool in evaluating the allowable minimum size of the spectral subset for different applications and analytes of interest.<sup>45</sup>

Our results, showing the improvement of prediction accuracy by performing nonlinear calibration, are consistent with similar results from NIR absorption studies reported earlier.<sup>39,46</sup> The curved effects in our spectral data set can probably be attributed to the fluctuations in sampling volume due to the change in tissue phantom turbidity (absorption and scattering). In a seminal study on the introduction of nonlinearity on vibrational spectra, Wülfert and coworkers showed that the temperature-induced spectral variations may produce a change in peak area, peak width, or/and a spectral shift.<sup>47</sup> Since sample turbidity can cause significant changes in peak area (intensity scaling) and smaller distortions of the intrinsic width of the spectral bands (via overlap of specific absorption and Raman features), it is likely that such changes will introduce curved effects that cannot be adequately modeled by linear multivariate calibration schemes. We have recently investigated this phenomenon in Raman spectroscopy for typically observed ranges of tissue absorption and scattering and obtained similar results.<sup>48</sup> The improved performance of SVR can also be attributed to the assessment (weighting) of the calibration samples by means of Lagrange

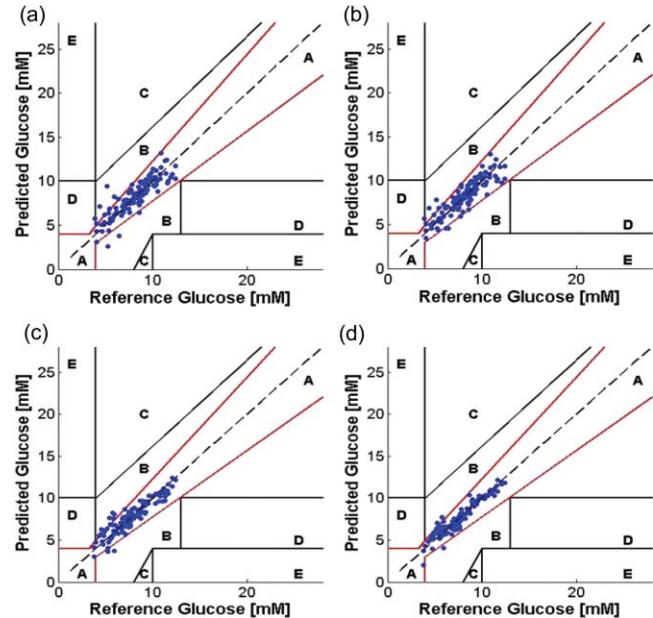
multipliers, which facilitates the ability to discriminate between important and irrelevant samples.<sup>39,49,50</sup>

Finally, we employed horizontal binning of the pixels to see if the adjacent pixel intensities could be combined without substantially compromising predictive ability. In principle, the numerical binning of pixel intensities of the Raman spectra provides a trade-off between higher spectral resolution and higher system throughput (analogous to increasing the slit width of a spectrograph). Here, our motivation is to investigate if it is possible to sample an even smaller set of spectral points than that prescribed by the wavelength selection approach above. To this end, we binned successive pixels using the selected 300 spectral points for SVR models with increasing bin sizes of one to six pixels (i.e., a total of 300 to 50 spectral points sampled). We observed that an increase in bin size to two pixels slightly reduces RMSEP from 0.63 to 0.59 mM. This can be attributed to the reduction in influence of spectrograph drift as well as the increase in SNR of each data point. Indeed, a bin size of three pixels provided an error value of 0.7 mM, comparable to that obtained from no numerical binning. This is borne out by the lack of statistically significant differences ( $p$ -value = 0.12) between the no binning and three pixel binning cases. However, further increase in bin size to four, five, and six pixels exhibits statistically significant increases in RMSEP to 0.84, 1.1, and 1.4 mM, respectively ( $p$ -value < 0.05 for each successive case). Our results suggest that combined use of wavelength selection and numerical binning can reduce the number of sampled wavelengths from 1000 to approximately 100 spectral points without a corresponding reduction in prediction accuracy.

#### 4.2 Human Subject Study

From Sec. 4.1, we observe that reduction of spectral points (even by a factor of three) does not substantially deteriorate calibration model performance, as long as the appropriate wavelengths depending on the analyte of interest and spectral interferences are analyzed. Nevertheless, the prospective transfer of the selected wavelength subset from one human to another is substantially more difficult because of the complexity arising from substantive variations in tissue optical (e.g., turbidity, autofluorescence, and skin heterogeneity) and physiological properties (dynamics of the analyte of interest) as well as changes in experimental conditions. Indeed, while several investigators have successfully applied wavelength selection in powder and mixture samples, to the best of our knowledge, its transferability in complex biological specimens has not been previously demonstrated. As previously mentioned, our goal here is to quantify the transferability of the selected wavelength subsets in such biological specimens and not that of the calibration models, which are separately developed on the individual human subjects as detailed in Sec. 3.2.

Figure 3 shows the prediction results of PLS and SVR calibration models using wavelength selection on the human subject dataset. The prediction results are shown plotted on Clarke error grids, which are widely used for quantifying the clinical accuracy of blood glucose monitors.<sup>18</sup> Predictions in zones A and B are regarded as acceptable while those in zones C, D, and E are potentially dangerous if used for determining treatment options. The RMSEP and the  $R^2$  correlation-coefficient values in each



**Fig. 3** Glucose prediction performance of wavelength-selected PLS and SVR calibration models in human subjects shown on the Clarke error grid. PLS calibration results are shown in (a) and (b) for 300 and 900 spectral points respectively. SVR calibration results are shown in (c) and (d) for 300 and 900 spectral points respectively.

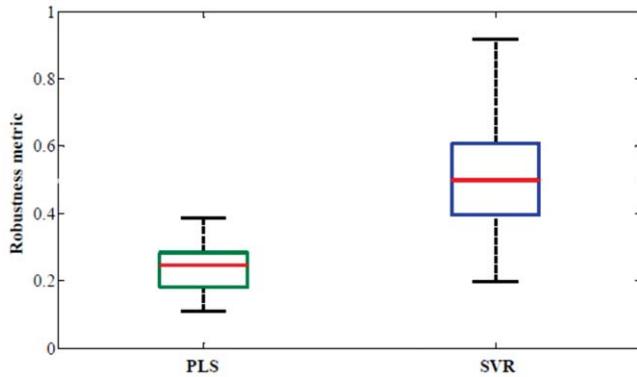
of the four cases (i.e., for PLS and SVR calibration with 300 and 900 spectral points, respectively) is provided in Table 3. We observe that the average RMSEP for SVR in each case is much lower than the corresponding value for PLS, such that the SVR calibration model with 300 spectral points provides equivalent (or better) prediction accuracy as compared to the PLS 900 spectral point model.

Current FDA recommendations (ISO 15197 guideline) stipulates that for clinical usage 95% of the sensor predictions should be within 15 mg/dL (0.83 mM) of reference for glucose <75 mg/dL (4.2 mM) and within 20% for glucose  $\geq$ 75 mg/dL (4.2 mM).<sup>51</sup> The percentage of data points satisfying the FDA criteria for our four calibration models are also tabulated in

**Table 3** Comparison of wavelength selection for PLS and SVR calibration models in human subjects.

	PLS		SVR	
	300 spectral points	900 spectral points	300 spectral points	900 spectral points
RMSEP	18.6	16.9	15.1	11.3
$R^2$	0.87	0.89	0.92	0.95
Percentage of points satisfying ISO criteria	86.66	89.17	88.33	94.17

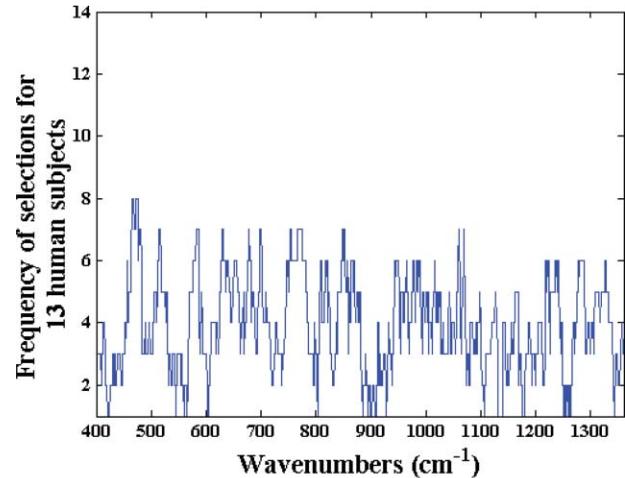
We report the enhanced robustness obtained for quantitative biological Raman spectroscopy by employing feature selection-based nonlinear support vector calibration. Importantly, we demonstrate the transferability of spectral subsets from one human subject to another for transcutaneous blood glucose measurements.



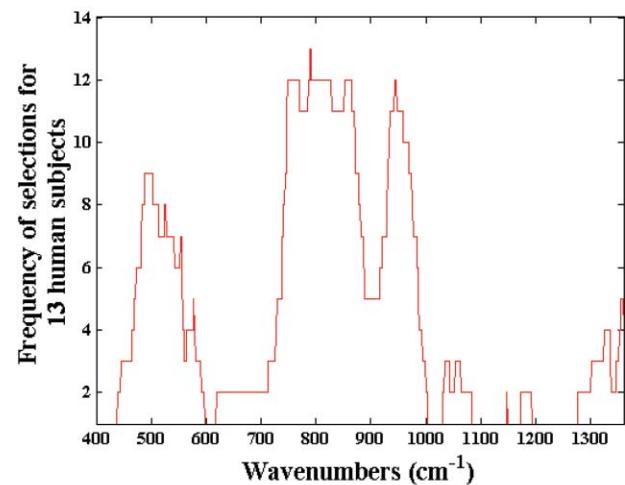
**Fig. 4** Boxplot showing robustness metric of 300 spectral points for PLS and SVR calibration models.

Table 3. While none of the models completely satisfy the 95% criterion (even though the SVR model for 900 spectral points is very close), one can reasonably expect that the deviations from the aforementioned criterion can be overcome by correcting for variations in tissue turbidity and autofluorescence as well as by addressing the physiological lag between blood and interstitial fluid glucose.<sup>16</sup> We expect that our current clinical studies, in collaboration with the MIT Clinical Research Center, encompassing nearly 100 normal and diabetic volunteers will provide the necessary datasets where we can validate our algorithms, including the application of wavelength selected support vector machines. Through these studies, we anticipate the demonstration of clinical feasibility of Raman spectroscopy for noninvasive blood glucose sensing.

Importantly, our results reflect that the selected wavelength subsets are transferable from one human subject to another. To quantify the transferability, we compare the degree of overlap between the wavelength subsets giving the minimum residue error for each of the human subjects. Evidently, if the degree of overlap between wavelength subsets from two human subjects is high, prospective application of the selected wavelength subset from one subject to the other will also provide accurate results. Here, we define robustness metric as the ratio of overlapping spectral points (between selected wavelength subsets of any two volunteers) to the total number of selected spectral points (300). Figure 4 provides box-plots of the robustness metrics for PLS and SVR calibration. The box-plots are generated from the robustness metrics evaluated for all possible combinations in the human subject dataset (i.e., total of 78 data-points from 13 subjects). Clearly, SVR provides a marked improvement with respect to transferability of the selected wavelength subsets. This information can also be readily visualized by a frequency plot showing the cumulative number of times a given spectral point is selected between the different models developed on individual human subjects. Figure 5 provides this alternate representation in which we observe the evident benefits of employing SVR in relation to PLS. The consistency of region selection in SVR is highlighted by the presence of greater structure (higher retention frequency of specific informative wavelengths) as compared to PLS modeling. Taken together, Figs. 4 and 5 illustrates the substantial enhancement of robustness provided by SVR modeling. This enhancement opens new avenues toward construction of universal calibration models based on a small set of features,



(a)



(b)

**Fig. 5** Frequency of wavelength selections for 13 human subjects using (a) PLS and (b) SVR calibration.

which also enables the development of Raman instruments employing a smaller set of wavelength sampling channels. A similar idea has been previously proposed by Buydens and coworkers,<sup>46</sup> arising from their observation of the robustness of SVR with respect to nonlinear effects in NIR absorption and Raman spectra, in regard to the usage of cheaper low-resolution spectroscopic systems for industrial applications. Our results here validate this line of thought in turbid biological media, where multiple interferences (i.e., other Raman scatterers, absorbers and fluorophores) exist.<sup>48,52</sup>

Finally, we note the implications of these results for the prospects of a miniaturized Raman instrument that can provide adequate detection sensitivity. As is well-known, a major drawback in translation of Raman spectroscopy is the large spatial footprint of the conventional Raman spectrometers, which are ill-suited to meet the needs of a clinical setting. Specifically, continuous glucose monitoring necessitates the development of a hand-held or a wearable device due to the frequency of measurements required for diabetic patients. To this end, Vo-Dinh and coworkers have explored, in a series of publications, the possibility of employing tunable detection filter-based (e.g.,

acousto-optic tunable filter) serial scanning Raman systems to significantly overcome the footprint drawbacks.<sup>19,20</sup> However, since the proposed systems employ serial acquisition of the photons at different wavelengths, a significant number of the Raman photons are not utilized in constructing the final spectrum and consequently larger acquisition times are needed to collect high quality Raman spectrum. While this problem cannot be completely eliminated, appropriate application of wavelength selection can greatly alleviate the problems associated with serial acquisition. For the specific problem of transcutaneous glucose detection, we have demonstrated above that a set of 300 spectral points can provide equivalent levels of accuracy as the full spectrum (900 spectral points). This implies that instead of having to perform serial acquisition over 900 spectral points, one can acquire the Raman photons at one-third of the spectrum, which in turn means the total acquisition time can be reduced by a factor of three. Conversely (and more importantly for glucose detection), one can acquire for longer periods of time (three-fold) at the appropriate wavelengths that leads to more efficient utilization of the important Raman photons. Considering shot-noise limited detection, a three-fold increase in acquisition time translates to a 1.73 times increase in analyte SNR. Using the minimum detectable concentration formulation,<sup>53,54</sup> this increase in SNR results in a corresponding reduction in the prediction uncertainty (or rise in precision) by nearly 43%.

## 5 Conclusion

In the present study, we have employed wavelength selection for linear (PLS) and nonlinear calibration (SVR) using tissue phantom and human subject datasets. We have demonstrated that the prediction accuracy is substantially improved by using SVR. In fact, our studies indicate that SVR models can provide the same prediction accuracy with a small fraction of the spectral information as used in PLS full spectrum analysis. Relative predictive determinant analysis has also been used to infer the size of the minimum allowable spectral subset that can provide calibration models of acceptable predictive quality. Furthermore, we show the prospective transferability of a selected wavelength subset across different human subjects while maintaining reasonably constant levels of prediction accuracy. It is observed that the SVR models, in particular, demonstrate surprising robustness and consistency in the selection of spectral bands. We believe that the resultant increase in accuracy and robustness, alongside the promise of enabling smaller and cheaper serial scanning Raman instruments, makes this an important step in the clinical translation of quantitative Raman spectroscopy. The approach proposed in this article, namely the combination of feature selection with nonlinear calibration schemes, is sufficiently broad to work for other disease diagnostics as well as for compositional analysis of pharmaceutical tablets, forensic analysis, and other process monitoring applications.

## Acknowledgments

The authors wish to thank the NIH National Center for Research Resources for their Grant No. P41-RR02594, at the MIT Laser Biomedical Research Center. I.B. acknowledges support of the Lester Wolfe Fellowship from the Laser Biomedical Research Center.

## References

1. C. J. Frank, D. C. B. Redd, T. S. Gansler, and R. L. McCreery, "Characterization of human breast biopsy specimens with near-IR Raman spectroscopy," *Anal. Chem.* **66**, 319–326 (1994).
2. A. S. Haka, K. E. Shafer-Peltier, M. Fitzmaurice, J. Crowe, R. R. Dasari, and M. S. Feld, "Diagnosing breast cancer by using Raman spectroscopy," *Proc. Natl. Acad. Sci. U.S.A.* **102**, 12371–12376 (2005).
3. J. T. Motz, M. Fitzmaurice, A. Miller, S. J. Gandhi, A. S. Haka, L. H. Galindo, R. R. Dasari, J. R. Kramer, and M. S. Feld, "In vivo Raman spectral pathology of human atherosclerosis and vulnerable plaque," *J. Biomed. Opt.* **11**, 021003 (2006).
4. A. M. K. Enejder, T. W. Koo, J. Oh, M. Hunter, S. Sasic, M. S. Feld, and G. L. Horowitz, "Blood analysis by Raman spectroscopy," *Opt. Lett.* **27**, 2004–2006 (2002).
5. J. Chaiken, W. Finney, P. E. Knudson, R. S. Weinstock, M. Khan, R. J. Bussjager, D. Hagrman, P. Hagrman, Y. W. Zhao, C. M. Peterson, and K. Peterson, "Effect of hemoglobin concentration variation on the accuracy and precision of glucose analysis using tissue modulated, noninvasive, *in vivo* Raman spectroscopy of human blood: a small clinical study," *J. Biomed. Opt.* **10**, 031111 (2005).
6. H. Martens and T. Naes, *Multivariate Calibration*, Wiley, New York (1989).
7. D. M. Haaland and E. V. Thomas, "Partial Least-Squares Methods for Spectral Analyses, I. Relation to other quantitative calibration methods and the extraction of qualitative information," *Anal. Chem.* **60**, 1193–1202 (1988).
8. J. H. Kalivas, N. Roberts, and J. M. Sutter, "Global optimization by simulated annealing with wavelength selection for ultraviolet-visible spectrophotometry," *Anal. Chem.* **61**, 2024–2030 (1989).
9. R. Leardi, "Application of genetic algorithm-PLS for feature selection in spectral data sets," *J. Chemom.* **14**, 643–655 (2000).
10. Y. Liang, Y. Xie, and R. Yu, "Accuracy criteria and optimal wavelength selection for multicomponent spectrophotometric determinations," *Anal. Chim. Acta* **222**, 347–357 (1989).
11. Q. Ding, G. W. Small, and M. A. Arnold, "Genetic algorithm-based wavelength selection for the near-infrared determination of glucose in biological matrices: initialization strategies and effects of spectral resolution," *Anal. Chem.* **70**, 4472–4479 (1998).
12. E. B. Hanlon, R. Manoharan, T. W. Koo, K. E. Shafer, J. T. Motz, M. Fitzmaurice, J. R. Kramer, I. Itzkan, R. R. Dasari, and M. S. Feld, "Prospects for *in vivo* Raman spectroscopy," *Phys. Med. Biol.* **45**, R1–R59 (2000).
13. N. C. Dingari, I. Barman, G. P. Singh, J. W. Kang, R. R. Dasari, and M. S. Feld, "Investigation of the specificity of Raman spectroscopy in non-invasive blood glucose measurements," *Anal. Bioanal. Chem.* **400**, 2871–2880 (2011).
14. I. Barman, G. P. Singh, R. R. Dasari, and M. S. Feld, "Turbidity-corrected Raman spectroscopy for blood analyte detection," *Anal. Chem.* **81**, 4233–4240 (2009).
15. I. Barman, C. Kong, G. P. Singh, and R. R. Dasari, "Effect of photobleaching on calibration model development in biological Raman spectroscopy," *J. Biomed. Opt.* **16**, 011004 (2011).
16. I. Barman, C. Kong, G. P. Singh, R. R. Dasari, and M. S. Feld, "Accurate spectroscopic calibration for noninvasive glucose monitoring by modeling the physiological glucose dynamics," *Anal. Chem.* **82**, 6104–6114 (2010).
17. K. L. Bechtel, W. C. Shih, and M. S. Feld, "Intrinsic Raman spectroscopy for quantitative biological spectroscopy part II: experimental applications," *Opt. Express* **16**, 12737–12745 (2008).
18. A. M. K. Enejder, T. G. Seccina, J. Oh, M. Hunter, W. Shih, S. Sasic, G. L. Horowitz, and M. S. Feld, "Raman spectroscopy for noninvasive glucose measurements," *J. Biomed. Opt.* **10**, 031114 (2005).
19. B. M. Cullum, J. Mobley, Z. Chi, D. L. Stokes, G. H. Miller, and T. Vo-Dinh, "Development of a compact handheld Raman instrument with no-moving parts for use in field analysis," *Rev. Sci. Instrum.* **71**, 1602–1607 (2000).
20. J. Mobley, B. M. Cullum, A. L. Wintenberg, S. S. Frank, R. A. Maples, D. L. Stokes, and T. Vo-Dinh, "Single-board computer based control system for a portable Raman device with integrated chemical identification," *Rev. Sci. Instrum.* **75**, 2016–2023 (2004).
21. L. Xu and I. Schechter, "Optimization method for simultaneous kinetic analysis," *Anal. Chem.* **68**, 2392–2400 (1996).

22. C. H. Spiegelman, M. J. McShane, M. J. Goetz, M. Motamedi, Q. L. Yue, and G. L. Cote, "Theoretical justification of wavelength selection in PLS calibration: development of a new algorithm," *Anal. Chem.* **70**, 35–44 (1998).
23. D. Jouan-Rimbaud, B. Walczak, D. L. Massart, I. R. Last, and K. A. Prebble, "Comparison of multivariate methods based on latent vectors and methods based on wavelength selection for the analysis of NIR spectroscopic data," *Anal. Chim. Acta* **304**, 285–295 (1995).
24. M. Zeaiter, J. M. Roger, and V. Bellon-Maurel, "Robustness of models developed by multivariate calibration. Part II: The influence of pre-processing methods," *Trends Analyt. Chem.* **24**, 437–445 (2005).
25. I. P. Soares, T. F. Rezende, R. C. Silva, E. V. R. Castro, and I. C. P. Fortes, "Multivariate calibration by variable selection for blends of raw Soybean Oil/Biodiesel from different sources using fourier transform infrared spectroscopy (FTIR) spectra data," *Energy Fuels* **22**, 2079–2083 (2008).
26. B. Igne and C. R. Hurburgh, Jr., "Local chemometrics for samples and variables: optimizing calibration and standardization processes," *J. Chemom.* **24**, 75–86 (2010).
27. F. Lindgren, P. Geladi, S. Rannar, and S. Wold, "Interactive variable selection (IVS) for PLS. Part 1: Theory and algorithms," *J. Chemom.* **8**, 349–363 (1994).
28. M. Forina, C. Casolino and C. Pizarro Millan, "Iterative predictor weighting (IPW) PLS: a technique for the elimination of useless predictors in regression problems," *J. Chemom.* **13**, 165–184 (1999).
29. V. Centner, D. L. Massart, O. E. de Noord, S. de Jong, B. M. Vandeginste, and C. Sterna, "Elimination of uninformative variables for multivariate calibration," *Anal. Chem.* **68**, 3851–3858 (1996).
30. R. Leardi and L. Norgaard, "Sequential application of backward interval partial least squares and genetic algorithms for the selection of relevant spectral regions," *J. Chemom.* **18**, 486–497 (2004).
31. J. Jiang, R. J. Berry, H. W. Siesler, and Y. Ozaki, "Wavelength interval selection in multicomponent spectral analysis by moving window partial least-squares regression with applications to mid-infrared and near-infrared spectroscopic data," *Anal. Chem.* **74**, 3555–3565 (2002).
32. S. Kasemsumran, Y. P. Du, K. Murayama, M. Huehne, and Y. Ozaki, "Near-infrared spectroscopic determination of human serum albumin,  $\gamma$ -globulin, and glucose in a control serum solution with searching combination moving window partial least squares," *Anal. Chim. Acta* **512**, 223–230 (2004).
33. Y. P. Du, Y. Z. Liang, J. H. Jiang, R. J. Berry, and Y. Ozaki, "Spectral regions selection to improve prediction ability of PLS models by changeable size moving window partial least squares and searching combination moving window partial least squares," *Anal. Chim. Acta* **501**, 183–191 (2004).
34. S. Sahin, E. Sarburun, and C. Demir, "Net analyte signal-based simultaneous determination of dyes in environmental samples using moving window partial least squares regression with UV-vis spectroscopy," *Analytical Methods* **1**, 208–214 (2009).
35. V. Vapnik, *The Nature of Statistical Learning Theory*, Springer-Verlag, Berlin (1995).
36. S. Ramaswamy, P. Tamayo, R. Rifkin, S. Mukherjee, C. Yeang, M. Angelo, C. Ladd, M. Reich, E. Latulippe, J. P. Mesirov, T. Poggio, W. Gerald, M. Loda, E. S. Lander, and T. R. Golub, "Multiclass cancer diagnosis using tumor gene expression signatures," *Proc. Natl. Acad. Sci. U.S.A.* **98**, 15149–15154 (2001).
37. S. Zomer, R. G. Brereton, J. F. Cartera, and C. Eckersb, "Support vector machines for the discrimination of analytical chemical data: application to the determination of tablet production by pyrolysis-gas chromatography-mass spectrometry," *Analyst* **129**, 175–181 (2004).
38. B. Scholkopf and A. J. Smola, *Learning with Kernels*, MIT Press, Cambridge (2002).
39. U. Thissen, B. Üstün, W. J. Melssen, and L. M. C. Buydens, "Multivariate Calibration with Least-Squares Support Vector Machines," *Anal. Chem.* **76**, 3099–3105 (2004).
40. A. Borin, M. F. Ferrao, C. Mello, D. A. Maretto, and R. J. Poppi, "Least-squares support vector machines and near infrared spectroscopy for quantification of common adulterants in powdered milk," *Anal. Chim. Acta* **579**, 25–32 (2006).
41. N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines: and Other Kernel-based Learning Methods*, Cambridge University, Cambridge (2000).
42. W. C. Shih, "Quantitative biological Raman spectroscopy for non-invasive blood analysis," Massachusetts Institute of Technology, Mechanical Engineering, Ph.D. Thesis, Cambridge, Massachusetts (2007).
43. H. Mark, "Use of Mahalanobis distances to evaluate sample preparation methods for near-infrared reflectance analysis," *Anal. Chem.* **59**, 790–795 (1987).
44. C. C. Chang and C.-J. Lin, *LIBSVM: a library for support vector machines*, <http://www.csie.ntu.edu.tw/~cjlin/libsvm/> (2001).
45. P. Williams, *Near-infrared Technology in the Agricultural and Food Industries*, 2nd ed., P. Williams and K. Norris, Eds., AACC Inc., St. Paul, Minnesota (2001).
46. U. Thissen, M. Pepers, B. Üstün, W. J. Melssen, and L. M. C. Buydens, "Comparing support vector machines to PLS for spectral regression applications," *Chemom. Intell. Lab. Syst.* **73**, 169–179 (2004).
47. F. Wülfert, W. T. Kok, and A. K. Smilde, "Influence of Temperature on Vibrational Spectra and Consequences for the Predictive Ability of Multivariate Models," *Anal. Chem.* **70**, 1761–1767 (1998).
48. I. Barman, C. Kong, N. C. Dingari, R. R. Dasari, and M. S. Feld, "Development of Robust calibration models using support vector machines for spectroscopic monitoring of blood glucose," *Anal. Chem.* **82**, 9719–9726 (2010).
49. X. Shao, X. Bian, J. Liu, M. Zhang, and W. Cai, "Multivariate calibration methods in near infrared spectroscopic analysis," *Analytical Methods* **2**, 1662–1666 (2010).
50. R. G. Brereton and G. R. Lloyd, "Support vector machines for classification and regression," *Analyst* **135**, 230–267 (2010).
51. G. J. Kost, N. K. Tran, V. J. Abad and R. F. Louie, "Evaluation of point-of-care glucose testing accuracy using locally-smoothed median absolute difference curves," *Clin. Chim. Acta* **389**, 31–39 (2008).
52. D. Wu, Y. He, S. Feng, and D. Sun, "Study on Infrared Spectroscopy Technique for Fast Measurement of Protein Content in Milk Powder Based on LS-SVM," *J. Food Eng.* **84**, 124–131 (2008).
53. A. Lorber and B. Kowalski, "Estimation of prediction error for multivariate calibration," *J. Chemom.* **2**, 93–109 (1988).
54. O. R. Scepanovic, K. L. Bechtel, A. S. Haka, W. Shih, T. Koo, A. J. Berger, and M. S. Feld, "Determination of uncertainty in parameters extracted from single spectroscopic measurements," *J. Biomed. Opt.* **12**, 064012 (2007).