Least-squares support vector machines modelization for time-resolved spectroscopy

Chauchard, F; Roussel, S; Roger, JM; Bellon-Maurel, V; Abrahamsson, Christoffer; Svensson, Tomas; Andersson-Engels, Stefan; Svanberg, Sune

Published in:
Applied Optics

2005

Least-squares support vector machines modelization for time-resolved spectroscopy

Fabien Chauchard, Sylvie Roussel, Jean-Michel Roger, Véronique Bellon-Maurel, Christoffer Abrahamsson, Tomas Svensson, Stefan Andersson-Engels, and Sune Svanberg

By use of time-resolved spectroscopy it is possible to separate light scattering effects from chemical absorption effects in samples. In the study of propagation of short light pulses in turbid samples the reduced scattering coefficient and the absorption coefficient are usually obtained by fitting diffusion or Monte Carlo models to the measured data by use of numerical optimization techniques. In this study we propose a prediction model obtained with a semiparametric modeling technique: the least-squares support vector machines. The main advantage of this technique is that it uses theoretical time dispersion curves during the calibration step. Predictions can then be performed by use of data measured on different kinds of sample, such as apples.

OCIS codes: 300.6500, 000.3860, 170.3660, 290.7050.

1. Introduction

Striking advances have been made in time-resolved spectroscopy (TRS).[1] Whereas conventional near-infrared spectroscopic measurements are influenced by light scattering in the sample, TRS deconvolutes absorption from scattering effects. The scattering properties of a sample depend on the physical properties of the samples, whereas the absorption is mostly dependent on the chemical composition of the samples. TRS was first developed for medical applications[2,3] but is now extended to other fields, such as pharmaceutical applications[4] and agricultural applications.[5,6] TRS uses short laser pulses of a few picoseconds to irradiate a sample. The light signal diffusively remitted by the sample at a given distance from the irradiation point is then temporally recorded.[7] The recordings can be made in either the reflection or the transmission mode. To measure the temporal signal at different wavelengths simultaneously, new techniques that use a streak camera for detection have been proposed. There are different ways to obtain light pulses with a broad wavelength profile; one is to use continuum generation by focusing a high power laser pulse in a cuvette of water.[8] The development of photonic crystal fibers has further simplified the instrumental setups for continuum generation.[9]

Once the two-dimensional signal is recorded, the reduced scattering coefficient (μ_s') and the absorption coefficient (μ_a) are obtained by linking the experimental data with theoretical or modeling data. This step is crucial to obtain the correct results, and many methods have been proposed. Three approaches are usually found: Monte Carlo simulations,[7,10] numerical optimizations,[11,12] and analytical descriptors of temporal dispersion.[13] Since the signal cannot be described by a linear equation, a nonlinear multivariate model is required. Semiparametric methods, such as kernel methods, provide more understandable models than artificial neural networks. Recently least-squares support vector machines (LSSVM)[14] methods have been developed and applied to near-infrared spectroscopy issues such as nonlinear discrimination[15,16] and quantitative predictions.[17]

Our aim here is to study LS SVM models calibrated only by use of theoretical data calculated from the diffusion equation in the reflectance mode. These models are then applied to predict the reduced scattering coefficient and the absorption coefficient of the experimental data.

F. Chauchard (fabien.chauchard@montpellier.cemagref.fr), J.-M. Roger, and V. Bellon-Maurel are with Information and Technologies for Agro-processes Cemagref, BP 5095, 34033 Montpellier, Cedex 1, France. S. Roussel is with Agrometrix, MINEA research parc, Cemagref, 361 rue Jean-Francois Breton, BP 5095, 34033 Montpellier, Cedex 1, France. C. Abrahamsson, T. Svensson, S. Andersson-Engels, and S. Svanberg are with the Department of Physics, Lund Institute of Technology, P.O. Box 118, Lund SE-221 00, Sweden.

Received 28 February 2005; revised manuscript received 10 May 2005; accepted 16 May 2005.

0003-6935/05/337091-07$15.00/0

2. Theory

A. Diffusion Equation

Photon transport in turbid media is described by the radiative transport equation[18]:

$$\frac{1}{c}\frac{\partial L(r,\ \mathbf{s},\ t)}{\partial t} + \mathbf{s}\cdot\nabla L\ (r,\ \mathbf{s},\ t) + (\mu_s + \mu_a)\, L(r,\ \mathbf{s},\ t)$$

$$= \mu_s \int_{4\pi} L(r,\ \mathbf{s},\ t) p(\mathbf{s},\ \mathbf{s}')\mathrm{d}\omega' + Q(r,\ \mathbf{s},\ t). \quad (1)$$

Here L is the radiance at a given distance r from the irradiating source at time t and in direction $\mathbf{s}$; $p(\mathbf{s}, \mathbf{s}')$ is the Henyey–Greenstein phase function; $\mathrm{d}\omega'$ is the angle between the initial photon direction $\mathbf{s}$ and the new direction $\mathbf{s}'$; and c is the speed of light in vacuum. To solve this equation the sample geometry must be taken into account. In the case of a semi-infinite homogeneous medium measured in reflection, the solution is given by the time-resolved diffusion equation[7]

$$R(\rho,\ t) = (4\pi D\nu)^{-3/2} z_0 t^{-5/2} \exp(-\mu_a \nu t) \times \exp\left(-\frac{\rho^2 + {z_0}^2}{4D\nu t}\right). \quad (2)$$

Here R is the signal measured at a given distance ρ at time t; D is the diffusion coefficient with $D(\lambda) = \{3[\mu_a(\lambda) + \mu_s'(\lambda)]\}^{-1}$; $z_0(\lambda) = [1/\mu_s']$ is the mean path; and ν is the speed of light in the medium, assumed to be constant in the measured wavelength range. A theoretical database containing time-resolved curves can be easily obtained by use of Eq. (2). A model can be derived based on this database, which can then be used to predict μ_a and μ_s'.

B. Least-Squares Support Vector Machines Theory

LS SVM models constitute an alternate formulation of SVM regression[19] proposed by Suykens.[14] Whereas classical multivariate regression is built on variables (e.g., time data for TRS or wavelengths for spectroscopic data) LS SVM methods are based on a kernel matrix **K.** The raw data matrix $\mathbf{X}_{n,p}$ containing n samples with p variables (e.g., n time-resolved curves), is then replaced by the $\mathbf{K}_{n,n}$ kernel defined as

$$\mathbf{K} = \begin{pmatrix} k_{1,1} & \cdots & k_{1,n} \\ \vdots & \ddots & \vdots \\ k_{n,1} & \cdots & k_{n,n} \end{pmatrix}. \quad (3)$$

Here $k_{i,j}$ is given by the radial basis function

$$k_{i,j} = \exp\left(\frac{-\|\mathbf{x}_i^{\mathrm{T}} - \mathbf{x}_j^{\mathrm{T}}\|^2}{\sigma^2}\right), \quad (4)$$

and $\mathbf{x}_i^T$ is the time response for a TRS measurement. The variable space is hence replaced by a sample space of a high dimension where a sample is defined by its distance to the other samples contained in the database. The proper subspace for modeling is tuned with the σ^2 parameter. The higher the σ^2, the wider the Gaussian kernel. Put simply, $k_{i,j}$ represents the similarities between the $\mathbf{x}_i^T$ and the $\mathbf{x}_j^T$ time responses. The model equation is then

$$\hat{\mathbf{y}} = \mathbf{K}\boldsymbol{\beta} + \beta_0, \quad (5)$$

where $\hat{\mathbf{y}}$ is the predicted value, $\mathbf{K}$ is the kernel as defined by Eq. (3), $\boldsymbol{\beta}$ is the regression vector, and β_0 is the offset term. Furthermore, the LS SVM objective function takes into account the norm of the regression vector to increase the model robustness. The classical squared loss function is thus replaced by the following objective function:

$$\min(\mathbf{e}) = \min\left[\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{2} + \frac{1}{\gamma}\frac{(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\beta})}{2}\right], \quad (6)$$

where γ is a regularization parameter analogous to the regularization parameter of regularized artificial neural networks and is used to weigh β norm. Once σ^2 and γ are chosen, the model is trained after constructing the Lagrangian by solving the linear Karush–Kuhn–Tucker system:

$$\begin{bmatrix} 0 & \mathbf{l}_n^{\mathrm{T}} \\ \mathbf{l}_n & \mathbf{K} + \dfrac{\mathbf{I}}{\gamma} \end{bmatrix} \begin{bmatrix} \hat{b}_0 \\ \hat{\mathbf{b}} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix}, \quad (7)$$

where $\mathbf{I}$ refers to an $[n \times n]$ identity matrix and $\mathbf{l}_n$ is an $[n \times 1]$ unity vector. The solution of Eq. (7) can be found by use of most standard methods of solving sets of linear equations, such as the conjugate gradient descent.

3. Material and Methods

A. Instrumentation

Figure 1 depicts the experimental setup. The instrument has been described in detail elsewhere.[9] Briefly, a Ti:sapphire mode-locked laser, pumped by an Ar-ion laser, was used to generate 100 fs pulses centered around 800 nm with an 80 MHz repetition rate. The laser pulses were focused into a 100 cm long index guiding crystal fiber (ICF)(Crystal Fiber A/S, Copenhagen, Denmark). The broadband light pulses generated by nonlinear effects in the ICF ranged from 750 to 1100 nm. The light was then transferred by a set of lenses into a gradient-index fiber guiding the light to the sample. Another gradient fiber, with the distal tip 6 mm from the irradiating tip, was used to collect the light diffusively reflected by the sample. The fibers were put in contact with the sample. A streak camera (Hamamatsu Model C5680) coupled to an imaging spectrometer (Chromex Model 250IS) captured the reflected light as a function of time and wavelength $R(t, \lambda)$. The spectral resolution was 0.93

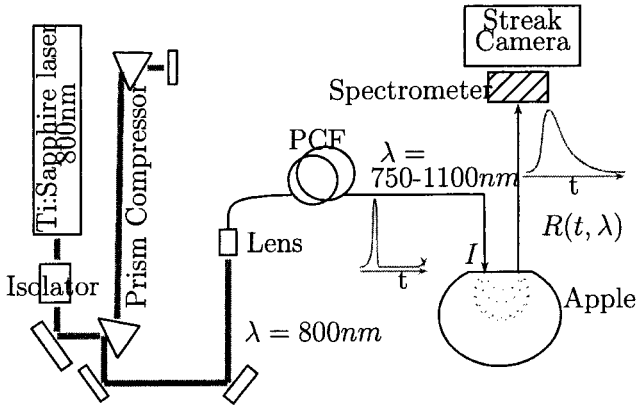


Fig. 1. Setup for TRS spectrum acquisition.

nm distributed over 512 pixels while the temporal resolution was 2.93 ps in the span from 0 to 1900 ps, spread over 640 pixels.

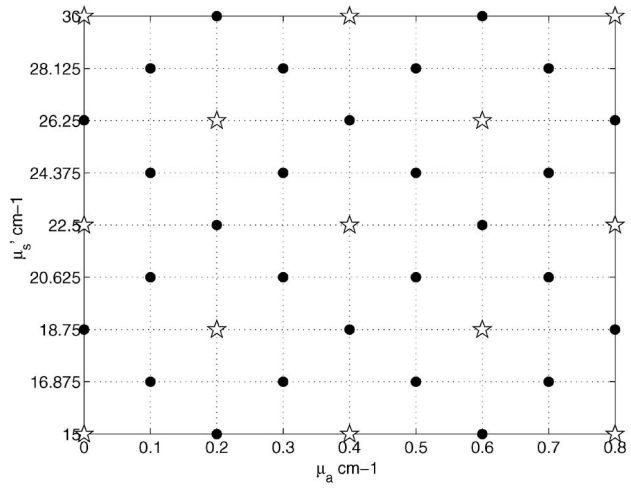
B. Measured Samples

Fifteen Golden Delicious apples were measured with the TRS setup. A small part of the apple was carefully removed to create a flat surface for applying the fibers. The measurements were performed immediately after the preparation of the apples to avoid flesh drying. Prior to each sample measurement, we recorded an instrumental response function by connecting the transmitting and receiving fibers to each end of a thin metal tube. This instrumental response function was used to determine time zero of the streak camera response and to measure the dispersion of the measured pulse that is due to the system characteristics.

C. Least-Squares Support Vector Machines Model

The LS SVM model was derived by use of a theoretical calibration set. We obtained the dataset by using the diffusion equation for an interfiber distance of $\rho = 6$ mm and a time resolution of 2.93 ps. Each signal was normalized by division with its maximum to become independent of variations in the irradiating signal intensity level. To improve the model efficiency, the temporal window between $t = 43$ and 900 ps was selected, where the time dispersion curves with different optical properties were significantly different. To span the absorption and scattering variations of apples,[20] a mixture design was used as described in Fig. 2. To tune γ and σ^2 the training set was split into two subsets, one for calibration (subset A) and one for validation (subset B). After the two parameters were chosen, the final model was constructed by use of the whole theoretical dataset.

The LS SVM toolbox (LS SVM version 1.4[21]) was used with MATLAB 6.0 (The MathWorks, Incorporated, Natick, Massachusetts) to derive the LS SVM models. To evaluate the accuracy of this new method, we compared the predicted values of μ_a and μ_s' with the ones fitted to the diffusion equation by using a Levenberg–Marquardt algorithm (LMA) previously used for apple TRS measurements.[20]



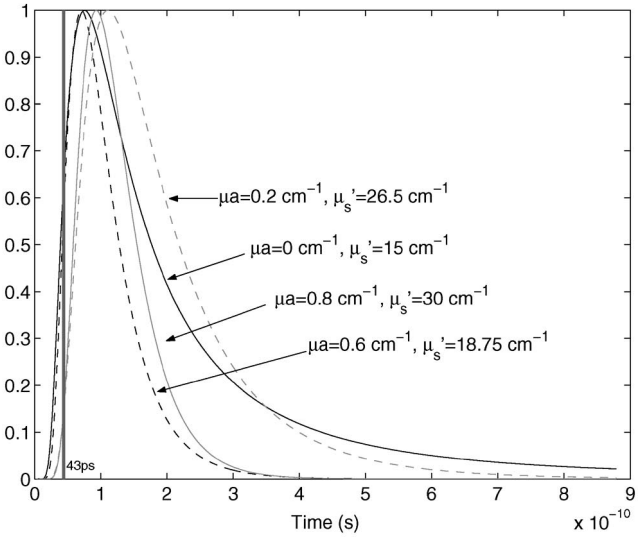


Fig. 2. (a) Values of μ_a and μ_s' for the training set: ★, calibration (A); ●, validation (B) for γ and σ^2 parameter tuning. (b) Four theoretical dispersion profiles.

4. Results and Discussion

A. Time-Resolved Spectroscopy Measurements

Figure 3(a) shows the instrumental response function. The continuum light pulses obtained were 300 nm wide (800–1100 nm). The temporal width was approximately 23 ps FWHM. The spectral profile was sensitive to changes in the laser intensity and variations in incoupling efficiency into the ICF. As a result, the spectral profile of the irradiating source changed from one sample measurement to another. The LS SVM model uses a temporal signal at a given wavelength. This response is normalized to obtain a maximum value of 1. Hence, source intensity variation from one sample measurement to another does not act on the model prediction efficiency. The recorded signal from one apple is depicted in Fig. 3(b). The temporal dispersion is high because of the scattering phenom-

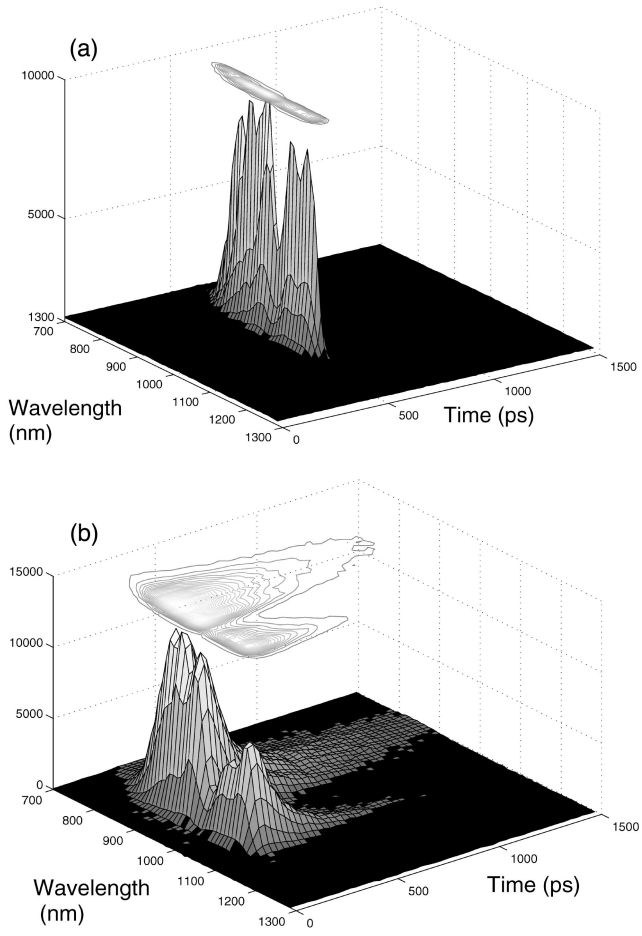


Fig. 3. Recording of two-dimensional time-resolved measurements: (a) multispectral light pulse for sample irradiation and (b) recorded signal for an apple.

ena inside the apple. Since the recorded signal-to-noise ratio was high enough in the region ranging from 800 to 1050 nm, this spectral window was selected for the study of optical properties of the apples.

B. Model Tuning

The optimization response surface for μ_a prediction is illustrated in Fig. 4. This surface represents the standard error of prediction on validation set B. The best prediction of μ_a was found for $\gamma = 50$ and $\sigma^2 = 500$. The μ_s' response surface (not presented here) gives an optimal solution for the same values. Since σ^2values are the same for both μ_a and μ_s', the kernel matrix is the same; this means that both models are built on the same subspace, allowing for the same degree of nonlinearities. Only the regression vectors are different for predicting μ_a and μ_s'. Low values of robustness criteria γ imply that regression vectors have a small norm that is necessary for a robust model.

C. Evaluation of Scattering and Absorption Coefficients for Experimental Data

Figure 5 shows a comparison of μ_a and μ_s' values predicted by the LMA and LS SVM for one apple. The absorption coefficient curves are similar, which

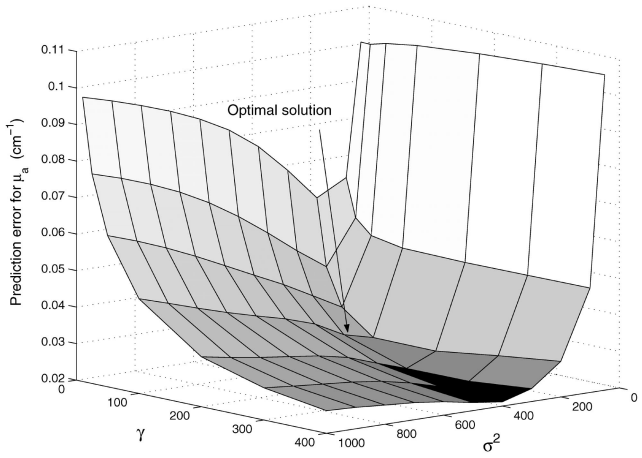


Fig. 4. Optimization surface for γ and σ^2 tuning for μ_a modelization.

proves the LS SVM prediction capabilities. In spite of the noise, the water peak is clearly visible at 970 nm as normally seen in conventional near-infrared spectra of fruits. With regard to the scattering coefficient the prediction values present an offset compared with the LMA results. This can be explained by the temporal dispersion shown in Fig. 6: the LS SVM model considers the irradiating peak as perfectly resolved in time (time width infinitely small), whereas the LMA takes the instrumental response into account in the calculations. Since the LMA is based on convolution, the predicted TRS curves are closer to the measured signal. However the LS SVM produce acceptable results. The LS SVM curves are above the LMA and are slightly peak shifted, which explains the offset previously observed between μ_s' values.

D. Prediction Performances

Figure 7 shows the LS SVM predicted values versus the LMA values of μ_a for the 15 apples (271 dispersion curves per sample). The determination coefficient of 0.96 is satisfactory, with a standard error of

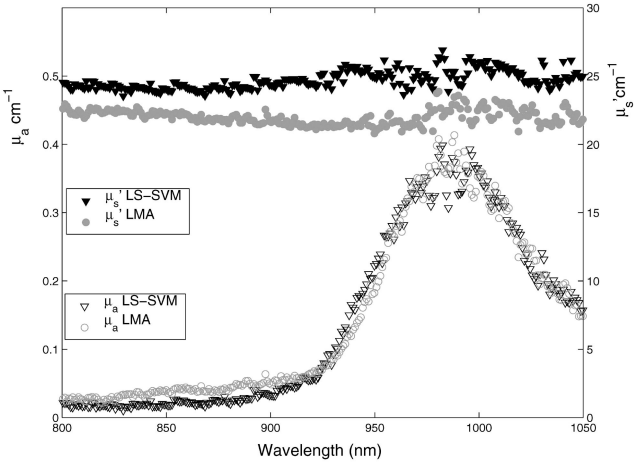


Fig. 5. Results for μ_a and μ_s prediction for an apple at all wavelengths.

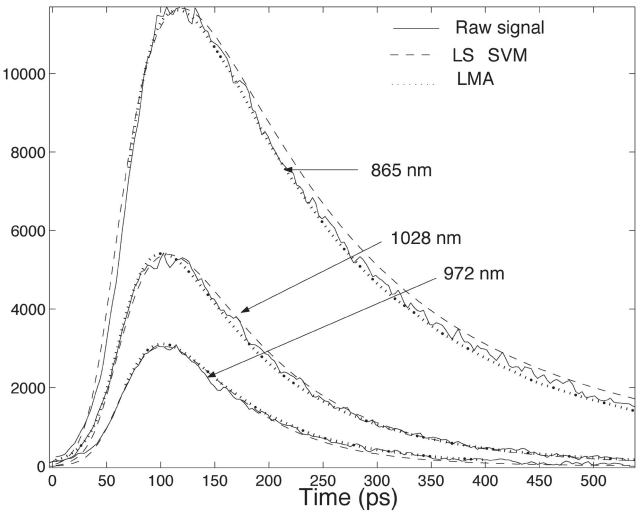


Fig. 6. Measured signal and fitted signals for three wavelengths.

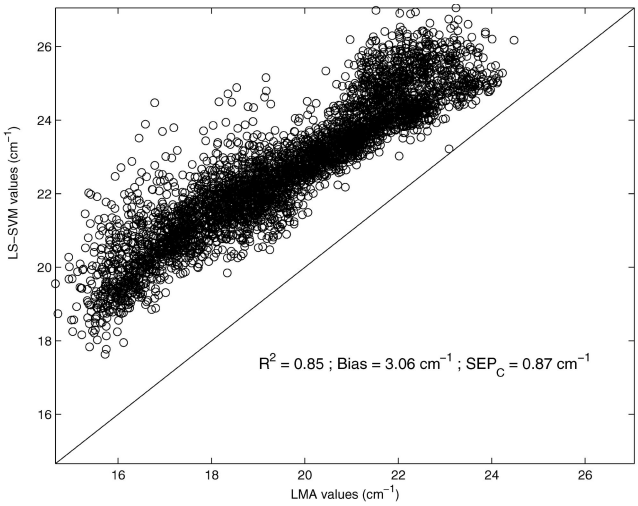


Fig. 8. Predicted performance of the μ_s' prediction model.

prediction of 0.02 cm^{-1}. It should be noted that there are no real reference values, but only reference values estimated by the LMA. Figure 8 shows a bias between LMA values and LS SVM predicted values for μ_s' determination. As explained in Subsection 4.C, this difference comes from the convolution process that is not used in the LS SVM. Since the determination coefficient is satisfactory (0.85), the model can easily be bias corrected by adding a constant (-3.06 cm^{-1}). However, this approach would consider LMA values as real reference values, although the LMA also has drawbacks and inaccuracies. For this reason, it would be more interesting to follow a more sophisticated approach, integrating a convolution process when building the database. In this case, the model would be calibrated on theoretical curves obtained by convoluting the diffusion equation with the instrumental response function. Of course, this method is more time-consuming since the model must be designed for each sample. When this approach is followed the prediction plot gives the results shown in Fig. 9. As assumed, the bias is reduced but is not small enough to be neglected. Furthermore, the correlation coefficient between LMA and LS SVM values decreases to 0.75. The noise in the measured data acts differently in the two methods since they have different bases. A visual curve analysis [Fig. 10(a)] is not accurate enough to be used to judge the differences between method performance. For this reason, the determination coefficient between the raw signal and the two estimated signals are presented for each wavelength in Fig. 10(b). The temporal curves calculated with the LS SVM predicted coefficient clearly have high performance ($r^2 > 0.99$), which is close to those calculated with the LMA. This tends to prove the accuracy of the proposed approach.

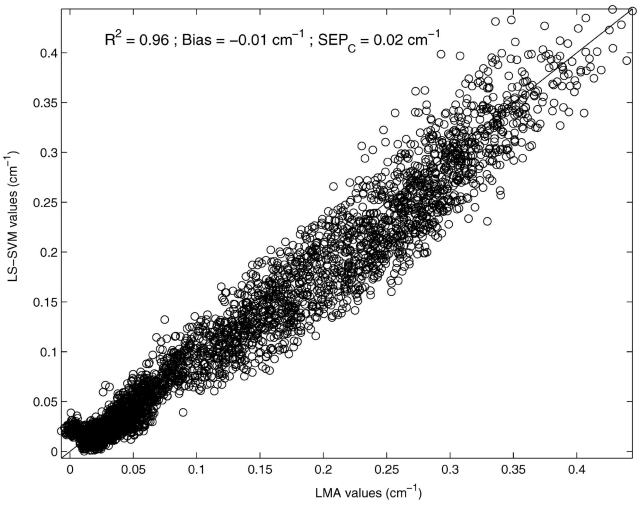


Fig. 7. Predicted performance of the μ_a prediction model.

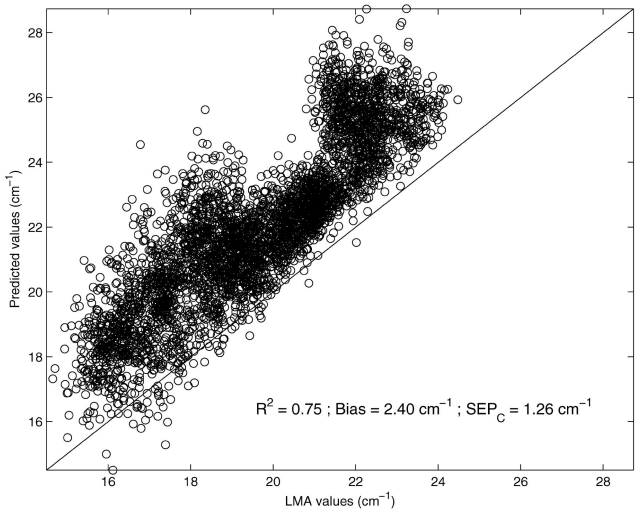


Fig. 9. Predicted performance of the μ_s' prediction model with a convolution approach.

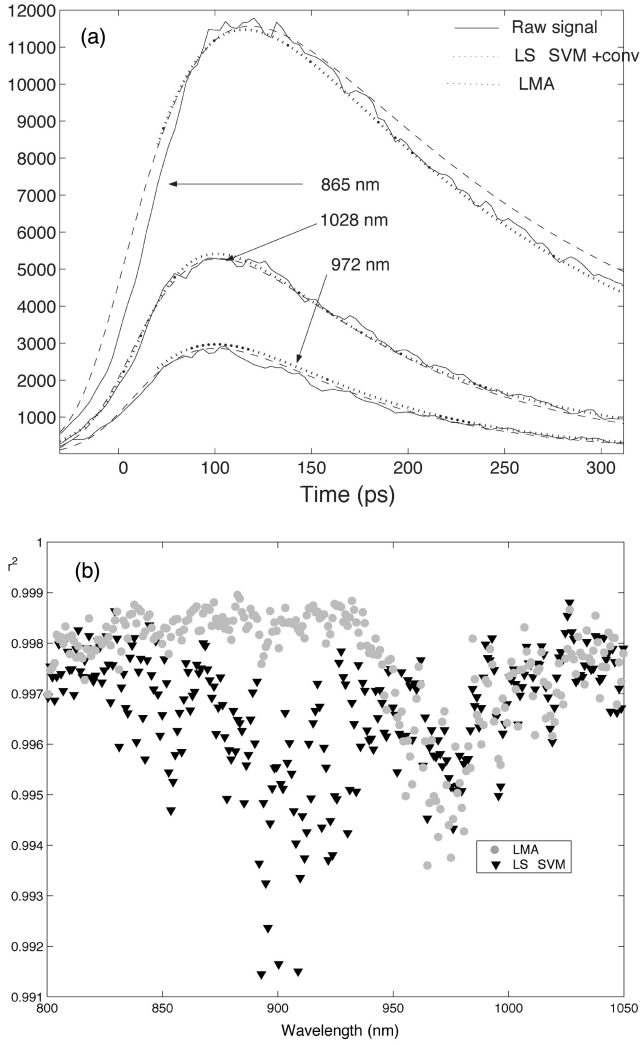


Fig. 10. Measured signal and fitted signals with the LS SVM obtained with convolution and determination coefficients: (a) predicted signals for three wavelengths and (b) r^2 between curves for each wavelength.

5. Conclusion

Thanks to its performance, the LS SVM model can be applied to time-resolved data for extraction of absorption and scattering coefficients. The model we have proposed in this paper has two main advantages. The first is that it can be used on any diffusing sample with $\mu_a < 0.08\ \mathrm{cm}^{-1}$ and $1.5\ \mathrm{cm}^{-1} < \mu_s' < 3\ \mathrm{cm}^{-1}$ (but a larger model can be calibrated) such as for human tissues. The second is that, since the model uses only 41 time-resolved curves, it can easily be integrated into an embedded sensor for industrial use. Even if the model performance is already interesting, the method can be improved by integration of a convolution process into the database construction. For optimization, data smoothing can be applied to the raw data.

The LS SVM could also be used with a database of Monte Carlo data. This would be interesting for measurement geometries for which the diffusion approximation is not valid, e.g., where the source and the detection fibers are situated close to each other or when the boundary conditions are too complex to be solved analytically.

As TRS transmission measurements produce the same type of curves as the reflection geometry, the LS SVM model can also be derived and applied efficiently to transmission data (slab geometry). We also believe that LS SVM modeling would be of great interest for spatially resolved spectroscopy and phase modulation spectroscopy.

This research was supported by the Integrated Initiative of Infrastructure project LASERLAB-EUROPE, contract RII3-CT-2003-506350.

References

1. B. Chance, J. Leigh, H. Miyake, D. Smith, S. Nioka, R. Greenfeld, M. Finander, K. Kaufmann, W. Levy, and M. Young, "Comparison of time-resolved and unresolved measurements of deoxyhemoglobin in brain," Proc. Natl. Acad. Sci. USA **85,** 4971–4975 (1988).
2. S. L. Jacques, "Time-resolved propagation of ultrashort laser pulses within turbid tissues," Appl. Opt. **28,** 2223–2229 (1989).
3. S. Andersson-Engels, R. Berg, S. Svanberg, and O. Jarlman, "Time-resolved transillumination for medical diagnostics," Opt. Lett **15,** 1179–1181 (1990).
4. J. Johansson, S. Folestad, M. Josefson, A. Sparen, C. Abrahamsson, S. Andersson-Engels, and S. Svanberg, "Time-resolved NIR/Vis spectroscopy for analysis of solids: pharmaceutical tablets," Appl. Spectrosc. **56,** 725–731 (2002).
5. P. E. Zerbini, M. Grassi, R. Cubeddu, A. Pifferi, and A. Torricelli, "Nondestructive detection of brown heart in pears by time-resolved reflectance spectroscopy," Postharvest Biol. Technol. **25,** 87–97 (2002).
6. J. Johansson, R. Berg, A. Pifferi, S. Svanberg, and L. Bjorn, "Time-resolved studies of light propagation in *Crassula* and *Phaseolus* leaves," Photochem Photobiol. **69,** 242–247 (1999).
7. M. S. Patterson, B. Chance, and B. C. Wilson, "Time-resolved reflectance and transmittance for the noninvasive measurement of tissue optical properties," Appl. Opt. **28,** 2331–2336 (1989).
8. S. Andersson-Engels, R. Berg, A. Persson, and S. Svanberg, "Multispectral tissue characterization with time-resolved detection of diffusely scattered white light," Opt. Lett. **18,** 1697–1699 (1993).
9. C. Abrahamsson, T. Svensson, S. Svanberg, S. Andersson-Engels, J. Johansson, and S. Folestad, "Time and wavelength resolved spectroscopy of turbid media using light continuum generated in a crystal fiber," Opt. Express **12,** 4103–4112 (2004).
10. T. J. Farrell, M. S. Patterson, and B. Wilson, "A diffusion theory model of spatially resolved, steady-state diffuse reflectance for the noninvasive determination of tissue optical properties *in vivo,*" Med. Phys. **19,** 879–888 (1992).
11. R. Cubeddu, A. Pifferi, P. Taroni, A. Torricelli, and G. Valentini, "Experimental test of theoretical models for time resolved reflectance," Med. Phys. **23,** 1625–1633 (1996).
12. S. J. Madsen, B. C. Wilson, M. S. Patterson, Y. D. Park, S. L. Jacques, and Y. Hefetz, "Experimental tests of a simple diffusion model for the estimation of scattering and absorption coefficients of turbid media from time resolved diffuse reflectance measurements," Appl. Opt. **31,** 3509–3517 (1992).
13. L. Leonardi and D. H. Burns, "Quantitative constituent measurements in scattering media from statistical analysis of photon time-of-flight distributions," Anal. Chim. Acta **348,** 543–551 (1997).
14. J. A. K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle, *Least Squares Support Vector Machines* (World Scientific, 2002).

15. A. I. Belousov, S. A. Verzakov, and J. von Frese, "Applicational aspects of support vector machines," J. Chemom. **16,** 482–489 (2002).
16. R. Goodacre, "Explanatory analysis of spectroscopic data using machine learning of simple, interpretable rules," Vib. Spectrosc. **32,** 33–45 (2003).
17. F. Chauchard, R. Cogdill, S. Roussel, J. M. Roger, and V. Bellon-Maurel, "Application of LS-SVM to non-linear phenomena in NIR spectroscopy: development of a robust and portable sensor for acidity prediction in grapes," Chemom. Intell. Lab. Syst. **71,** 141–150 (2004).
18. A. Ishimaru, *Wave Propagation and Scattering in Random Media* (Academic, 1978).
19. V. Vapnik and A. Lerner, "Pattern recognition using generalized portrait method," Autom. Remote Control **24,** 774–780 (1963).
20. R. Cubeddu, C. D'Andrea, A. Pifferi, P. Taroni, A. Torricelli, G. Valentini, C. Dover, D. Johnson, M. Ruiz-Altisent, and C. Valero, "Nondestructive quantification of chemical and physical properties of fruits by time-resolved reflectance spectroscopy in the wavelength range 650–1000 nm," Appl. Opt. **40,** 538–543 (2001).
21. LS SVMs toolbox, www.esat.kuleuven.ac.be/sista/lssumlab/.