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

F. Chauchard, S. Roussel, J.M. Roger, Véronique Bellon Maurel, C. Abrahamsson, T. Svensson, S. Andersson Engels, S. Svanberg

Least Squares-Support Vector Machines modelisation for Time Resolved Spectroscopy

F. Chauchard (1), S.Roussel(2),J.M. Roger(1), V. Bellon-Maurel(1),C. Abrahamsson(3),T. Svensson(3), S. Andersson-Engels(3), S. Svanberg(3).

1-Information and Technologies for Agro-processes, Cemagref BP 5095, 34033 Montpellier Cedex 1, France
2- Ondalys, www.ondalys.com
3- Physics, Lund Institute of Technology, P.O Box 118,SE-221 00 Sweeden

Abstract

Time Resolved Spectroscopy is able to separate the light scattering effect from the chemical absorption effect. This method is based on the time dispersion of light pulses into the scattering medium. The reduced scattering coefficient and the absorption coefficient are usually obtained using numerical optimization technique or Monte Carlo simulation. In this study, we propose to create a prediction model obtained using a semi-parametric modelisation method : the Least-Squares Support Vector Machine. The main advantage of this model is that it uses theorical curve of time dispersion during the calibration step. The prediction can then be performed on different kind of samples such as apples or biological tissue.

1 Introduction

Striking advances have been made in Time Resolved Spectroscopy (TRS).[1] While near infrared spectroscopy measurement are influenced by the light scattered by the sample, TRS separates chemical interactions from scattering interactions. TRS was firstly developed for medical applications[2,3] . It is now extended to other fields such as pharmaceutical[4] products or agricultural products[5,6] . TRS uses a laser pulse of a few pico-second to irradiate a sample. The light signal reflected by the sample at a given distance from the irradiation point is then temporally recorded in reflection or in transmission[7] . In order to measure simultaneously the temporal signal at different wavelengths, new techniques involving non-linear effects have been proposed. The first one used water self modulation for generating a white pulse.[8] The

latest technology is based on light continuum generation using a polycristal fiber.[9]

Once the 2-dimension signal is recorded, the reduced scattering coefficient μ_s' and the absorption coefficient μ_a are obtained by linking the experimental data with the theoretical diffusion equation. Probably because this step is the most crucial, many methods have been proposed to estimate this relationship. Three approaches are usually found : Monte Carlo simulation[7,10] , numerical optimization[11,12] , analytical decriptor of temporal dispersion[13] . Since the signal can not be described by a linear equation (this is why curves descriptors are often used), a non linear multivariate model is required.Semi-parametric methods, such as kernel ones provide more understandable modelisation than artificial neural network. Recently Least-Square Support-Vector Machines (LS-SVM)[14] methods have been developed and applied to near infrared spectrometry issues such as a non-linear discrimination[15,16] and quantitative predictions.[17]

This paper aims at studying LS-SVM theorical models calibrated only using the diffusion equation in reflectance mode. This model is then applied to the prediction of the reduced scattering coefficient and the absorption coefficient on experimental data.

2 Theory

2.1 Diffusion Equation

The photon response is described by the radiative transfer equation which is[18] :

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

Where L is the radiance at a given distance r from the irridiating source at time t and in direction $\mathbf{s}$. $p(\mathbf{s},\mathbf{s}')$ is the Henyey Greenstein phase function, $d\omega'$ is the angle between initial photon direction $\mathbf{s}$ and the new one $\mathbf{s}'$. c is the speed of ligth in vaccum. In order to solve this equation the sample geometry must be taken into account for generating hypothesis on boundaries. In the case of a semi-infinite homogeneous medium measured in reflection, the

solution is[7] :

$$R(\rho,t) = (4\pi Dv)^{-3/2} z_0 t^{-5/2} \exp(-\mu_a vt) \exp\left(-\frac{\rho^2 + {z_0}^2}{4Dvt}\right) \tag{2}$$

Where R is the signal measured at a given distance ρ at time t. D is the diffusion coefficient with $D(\lambda) = [3(\mu_a(\lambda) + \mu'_s(\lambda))]^{-1}$ and $z_0(\lambda) = \frac{1}{\mu'_s(\lambda)}$ is the mean path. v is speed of light in the medium, steady across wavelengths. Considering a constant distance, ρ, the equation dimension can be reduced :

$$R(t) = (4\pi Dv)^{-3/2} z_0 t^{-5/2} \exp(-\mu_a vt) \exp\left(-\frac{\rho^2 + {z_0}^2}{4Dvt}\right) \tag{3}$$

A theoretical database containing time-resolved curves may be easily obtained usig Eq. (3). A model can be derived based on this database to predict μ_a and μ'_s.

2.2 LS-SVM theory

LS-SVM models are an alternate formulation of SVM regression[19] proposed by Suykens.[14] Whereas classical mulivariate regression are built on variables (e.g. time data for TRS or wavelengths for spectroscopic data) LS-SVM methods are based on a kernel matrix $\mathbf{K}$. The 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 :

$$\mathbf{K} = \begin{pmatrix} k_{1,1} & \dots & k_{1,n} \\ \vdots & \ddots & \vdots \\ k_{n,1} & \dots & k_{n,n} \end{pmatrix} \tag{4}$$

Where $k_{i,j}$ is given by the RBF function :

$$k_{i,j} = e^{\frac{-\left\|\mathbf{x_i^T} - \mathbf{x_j^T}\right\|^2}{\sigma^2}} \tag{5}$$

Where $\mathbf{x_i^T}$ is the time response for a TRS measurement. The variable space is hence replaced by a sample space of a very high dimension where a sample is defined by its distance to the other samples contained in the database. The

proper subspace for modelisation is tuned with σ^2 parameter. The higher the σ^2 is, the wider is the gaussian kernel. Put simply, $k_{i,j}$ represents the similarities between $\mathbf{x_i^T}$ and $\mathbf{x_j^T}$ time responses. The model equation is then :

$$\hat{\mathbf{y}} = \mathbf{K}\boldsymbol{\beta} + \beta_0 \tag{6}$$

Where $\hat{\mathbf{y}}$ is the predicted value, $\mathbf{K}$ is the kernel as defined by Eq.4, $\boldsymbol{\beta}$ 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 in order to increase the model robustness. The classical squared loss function is thus replaced by the following objective function :

$$min(\mathbf{e}) = min\left[\frac{\sum_{i=1}^{n}(\mathbf{y_i} - \hat{\mathbf{y}}_\mathbf{i})^2}{2} + \frac{1}{\gamma}\frac{(\boldsymbol{\beta}^T\boldsymbol{\beta})}{2}\right] \tag{7}$$

γ is a regularization parameter analogous to the regularization parameter of regularized artificial neural networks, is used to weigh $\boldsymbol{\beta}$ norm. Once σ^2 and γ are chosen, the model is trained after constructing the Lagrangian by solving the linear Karush-Kuhn-Tucker (KKT) system :

$$\begin{bmatrix} 0 & \mathbf{l_n^T} \\ \mathbf{l_n} & \mathbf{K} + \frac{\mathbf{I}}{\gamma} \end{bmatrix} \begin{bmatrix} \hat{b}_0 \\ \hat{\mathbf{b}} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix} \tag{8}$$

Where $\mathbf{I}$ refers to an $[n \times n]$ identity matrix, $\mathbf{l}_n$is a $[n \times 1]$ unity vector. The solution of Eq. (8) can be found using most standard methods of solving sets of linear equations, such as conjugate gradient descent.

3 Material and methods

3.1 Instrumentation

Fig. (1) depicts the experimental setup. More details concerning the design made by the Lund Laser Center can be found in Abrahamsson and al.[9] . Ti:Sapphire Laser, mode-locked using an Ar-ion Laser, was used to generated $100fs$ pulses centered around 800nm with $80Mhz$ repetition rate. After optical treatments (isolator, prism compressor...), laser pulses were focused into a $100cm$ long index guiding crystal fiber (ICF)(Crystal fiber A/S, Copenhagen, Denmark). The light continuum generated by the non-linear effect ranged from $750nm$ till $1100nm$. It was then collimated through

an optical fiber aiming at irradiating the sample. Another fiber, situated at $6mm$ from the irradiating one, was used to pick up the light retrodiffused by the sample. The fibers were in contact with the sample. A Streak Camera (Hamamatsu, ModelC5680) coupled with an imaging spectrometer (Chromex, Model 250IS) captured the spectro-temporal signal $R(t, \lambda)$. The spectral resolution was $0.93nm$ for 512 pixels and temporal resolution was $2.93ps$ for the 640 pixels. A recorded signal was then composed of 640 time curves containing each 512 variables (640×512). The integration time was $5min$ ($300s$).

3.2 Measued samples

15 Golden Delicious Apples were measured using the TRS setup. A small part of the apple was carefully removed in order to create a flat surface for applying the fibers. The measurements were immediately performed in order to avoid flesh drying effect. Prior to each sample measurement, an instrumental response function was recorded by connecting two fiber to each end of a thin metal tube. This instrumental response function was then used to determine the time of the laser pulse in the streak camera and to measure the dispersion of the measured pulse due to the system characteristics. The sample measurement was corrected using this reference.

3.3 LS-SVM model

The LS-SVM model was derived using a theoretical calibration set obtained using the diffusion equation, applied for $\rho = 6mm$ and a time resolution of $2.93ps$. Each signal was divided by its maximum in order to get rid of irradiating signal intensity level. For improving the model efficiency, the temporal window $t = 43ps$ to $900ps$ was selected where curves were significantlly different. The curve start untill $t = 43ps$ was discard because there were no significant differences between curves. The temporal window ranging from $t = 43ps$ till $900ps$ was chosen for our modelisation. To span the absorption and scattering variations of apple[20] , a mixture design was set up as described by Fig. (2). In order to tune γ and σ^2 the training set was split in two subsets, one for calibration (subset A) and one for validation (subset B). Once the 2 parameters were chosen, the model was calibrated using the whole theoritical training set. The model was then tested on experimental

measurements.
In order to evaluate the accuracy of this new method, the predicted values of μ_a and μ'_s were compared to the ones calculated by Levenberg Marquart (LMA) optimization procedure, already used on apple TRS measurements.[20] The LS-SVM toolbox (LS-SVM v1.4 [1], Suykens, Leuven, Belgium) was used with MATLAB 6.0 (The MathWorks, Inc., Natick, USA) to derive all of the LS-SVM models.

4 Results and discussion

4.1 TRS measurements

Fig. (3-a) shows the instrumental function response. The continuum light pulse obtained using non-linear effect in the polycrystal fiber is of $300nm$ width ($800nm - 1000nm$). Temporal width is about $23ps$. The spectral profile was found very sensitive to changes of laser modelocking due to optics heating. A spectro-temporal recorded apple signal is described Fig. (3-b). Temporal dispersion is very high due to scattering phenomena inside the medium and reaches 1025ps. Since the continuum light ranges from 800 to $1050nm$, the same spectral window was selected for studying the optical properties of the sample.

4.2 Model tuning

The optimization responses surface for μ_a prediction is illustrated in Fig. (4). This surface represents the standard error of prediction (SEP) on the validation set B. The best prediction of μ_a where found for $\gamma = 50$ and $\sigma^2 = 500$. μ'_s response surface (not presented here) gives the optimal solution for the same values. Since σ^2 values 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 on non-linearities; only the regression vectors are different for predicting μ_a and μ'_s. Low values of robustness criteria, γ, imply the regression vectors have a small norm which is necessary for a robust model.

[1] www.esat.kuleuven.ac.be/sista/lssvmlab/

4.3 Evaluation of scattering and absorption coefficients on experimental data

Fig.5 compares μ_a and μ'_s values predicted by LMA and LS-SVM for a given apple. The absorption coefficient curves are very similar, which proves LS-SVM prediction capabilities. In spite of the noise, the water peak is clearly visible at $970nm$ as normally seen in classical NIR spectra of fruits. Regarding the scattering coefficient the prediction values present an offset compared to LMA results. This can be explained by the temporal dispersion Fig. (6) : LS-SVM model considers the irradiating peak as perfectly resolved in time (time width infinitely small), whereas LMA uses the instrumental response to convoluate the diffusion equation. As shown in Fig.6 , LMA curves better fit the experimental signal at the beginning of the curve ($t = 0$ to $t = 100ps$). Contrarly, the LS-SVM solution better fits the experimental curves from $t = 100ps$ until the end of the signal. While improving the prediction performance of μ'_s, the convolution process enhances the sensitivity of the method to the light source variations. Since it is well known that the end of the signal is strongly correlated to μ_a, absorption coefficient predicted by LS-SVM are very similar to LMA, despite the offset between μ'_s predictions.

4.4 Prediction performances

Fig. (7) shows the LS-SVM predicted values versus 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 prediction of $0.02cm^{-1}$. It should be noted that there are no real reference values, but only reference values estimated by LMA which might be somewhat unaccurate. Fig. (8) shows a bias between LMA values and LS-SVM predicted values for μ'_s determination. As explained before, this difference comes from the convolution process which is not used in LS-SVM. Since the determination coefficient is satisfactory (0.85), the model may be easly bias corrected by adding a constant ($-3.06cm^{-1}$). However, this approach would consider LMA values as real reference values, although LMA has also its drawbacks and unaccuracy. For this reason, it would be more interesting to follow a more sophisticated approach, integrating a convolution process in the database building. In this case, the model would be calibrated on theoretical curves obtained by convoluating the diffusion equation with the continuum pulse (instrumental function). Of course, this method is more timeconsumming since the model

must be designed for each sample. When this approach is followed the prediction plot gives the results shown on Fig.9. As assumed, the bias is reduced but is not small enough to be neglected. Futhermore, the correlation coefficient between LMA and LS-SVM values decreases to 0.75. However, the predicted LS-SVM time resolved signals (Fig.10) are closer to the raw signal than LMA solution. Actually, the noise of the measured data acts differently on the two methods since they are base on different working. Since visual curves analysis couldn't be accurate enough to judge differences between method performance, the determination coefficient between the raw signal and the two estimated signal was calculated for each wavelength Fig.11. The temporal curves calculated using LS-SVM predicted coefficient have clearly the highest performance (r^2 near 1). This tends to prove the accuracy of the proposed approach. This doesn't means LS-SVM outperform LMA since LMA may also be improved using different preprocessing and tuning.

5 Conclusion

Thanks to its performances, LS-SVM model can be applied to experimental data for absorption and scattering prediction. The model proposed in this paper has two main advantages. The first one is that it can be used on any diffusing sample with $\mu_a < 0.08cm^{-1}$ and $1.5cm^{-1} < \mu'_s < 3cm^{-1}$ (but a larger model may be calibrated) such as for human tissues. The second one is that since the model uses only 41 time resolved curve for the model, it can be easily integrated on an embedded sensor for industrial uses. Even if the model performances are already interesting, the method may be improved by integrating a convolution process in the database building. As mentionned previously, without outperforming LMA, the performances are higly improved. In the perspective of optimisation, a data smoothing may be applied on the raw data.
As TRS transmission measurements produce the same type of curves as the reflection geometry, LS-SVM model may also be derived and applied efficiently on transmission data (slab geometry). We also think LS-SVM modeling would be of deep interest for spatially resolved spectroscopy and phase modulation spectroscopy.

Acknowledgments

This work was supported by the Integrated Initiative of Infrastructure project LASERLAB-EUROPE, Contract No. RII3-CT-2003-506350. The authors also gratefully acknowledge Lund Laser Center team for their disponibility and kindness.

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. Jacques, "Time-resolved propagation of ultrashort laser pulses whithin turbid tissue," Appl. Optics **28**, 2331–2336 (1989).

[3] S. Andersson-Engels, R. Berg, A. Persson, S. Svanberg, and O. Jarlman, "Time-resolved transillumination for medical diagnostics," Opt.Lett **15**, 1179–1191 (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," Applied Spectroscopy **56**, 725–731 (2002).

[5] P. Zerbini, M. Grassi, R. Cubeddu, A. Pifferi, and A. Torricelli, "Nondestructive detection of brown heart in pears by time-resolved reflectance spectroscopy," Postharvest Biology and Technology **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," Photochemistry and Photobiology **62**, 242–247 (1999).

[7] M. Patterson, B. Chance, and B. Wilson, "Time-resolved reflectance and transmittance for the non invasive 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, and S. Andersson-Engels, "Time and wabvelength resolved spectroscopy of turbid media using light continuum generated in a crystal fiber," Optics Express **12**, 4103–4112 (2004).

[10] T. Farrell, M. S. Patterson, and B. Wilson, "A diffusion theory 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. Madsen, B. Wilson, M. Patterson, Y. Park, S. Jacques, and Y. Hefetz, "Experimental tests of a simple diffusion model for the estimation of scattering and absorption coefficients of turbid media from timeresolved diffuse reflectance measure," Appl. Opt. **31**, 3509–3517 (1992).

[13] L. Leonardi and D. H. Burns, "Quantitative constituent in scattering media from statistical analysis of photon time-of-flight distributions," Analytica Chimica 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 Publishing Co., Singapore, 2002).

[15] A. Belousov, S. Verzakov, and J. von Frese, "Applicational aspect of support vector machine," J. of Chemometrics **16**, 482–489 (2002).

[16] R. Goodacre, "Explanatory analysis of spectroscopic data using machine learning of simple, interpretable rules," Vibrational Spectroscopy **32**, 33–45 (2003).

[17] F. Chauchard, R. Cogdill, S. Roussel, J. 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," Chemometrics and Intelligent Laboratory Systems **71**, 141–150 (2004).

[18] A. Ishimaru, *Wave propagation and scattering in random media* (Academic Press, New York, 1978).

[19] V. Vapnik and A. Lerner, "Pattern recognition using generalized portrait method," Automation and Remote Control **24**, 774–780 (1963).

[20] R. Cubeddu, C. 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," Applied Optics **40**, 538–543 (2001).

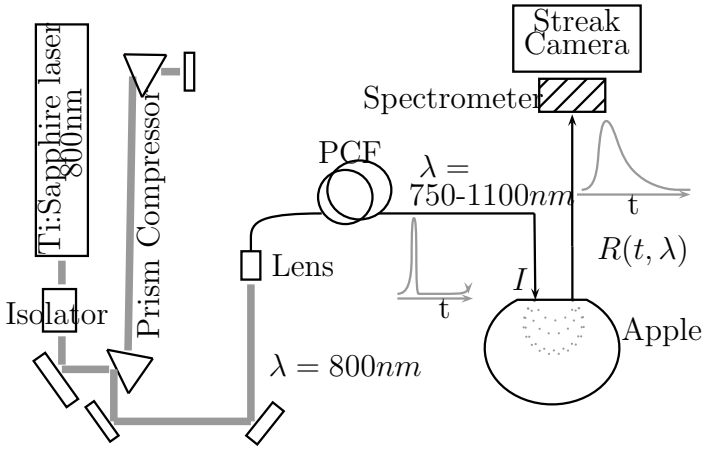
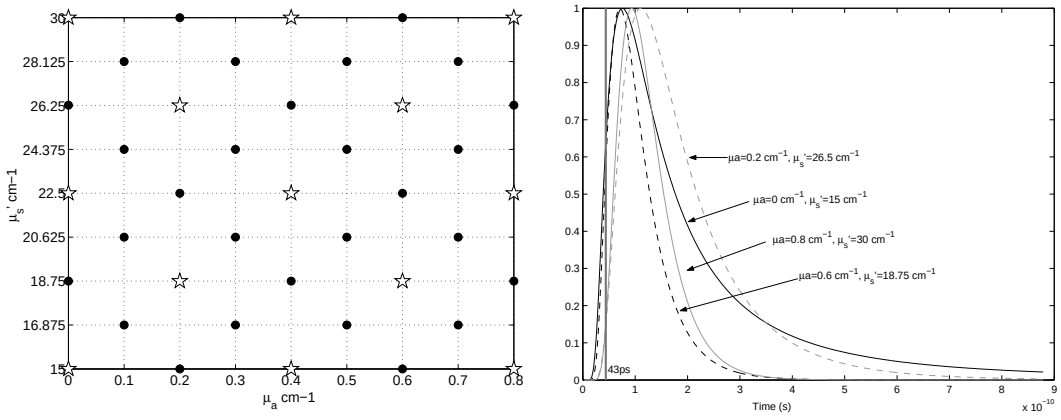


Figure 1: Setup for TRS spectrum acquisition



a- b-

Figure 2: (a)Values of μ_a and μ_s' for the training set ; $\star$ calibration(A), $\bullet$ validation(B) for γ and σ^2 parameters tuning. (b) Four theoritical dispersion profiles used are shown

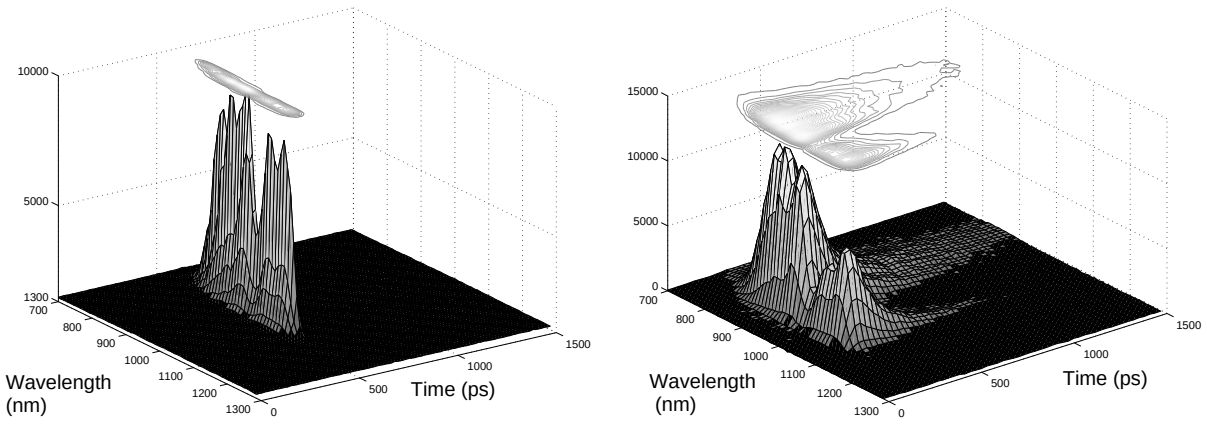


Figure 3: Recording of a 2D time resolved measurement : (a) Multispectral light pulse used for sample irradiation (b) Recorded signal for an apple

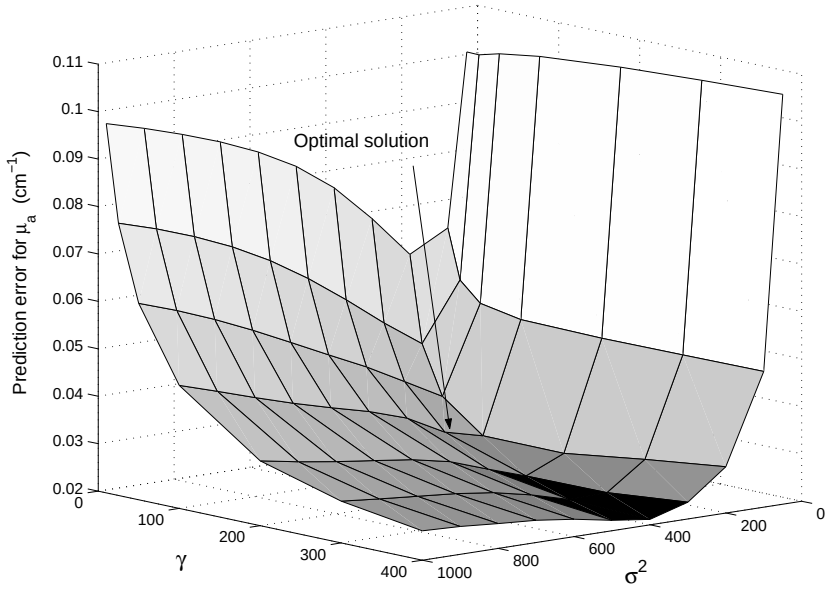


Figure 4: Optimization surface for γ and σ^2 tuning for μ_a modelisation

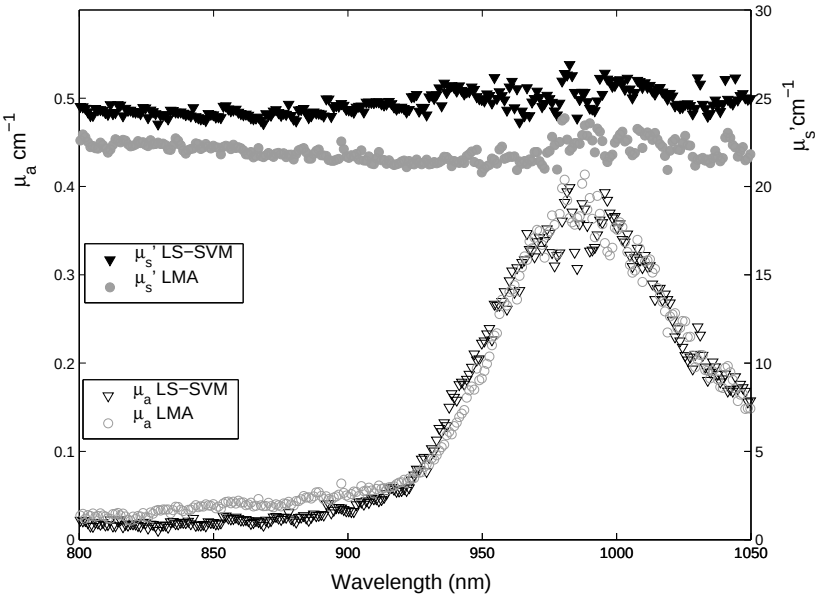


Figure 5: Result of μ_a and μ_s' prediction on an apple for all wavelengths

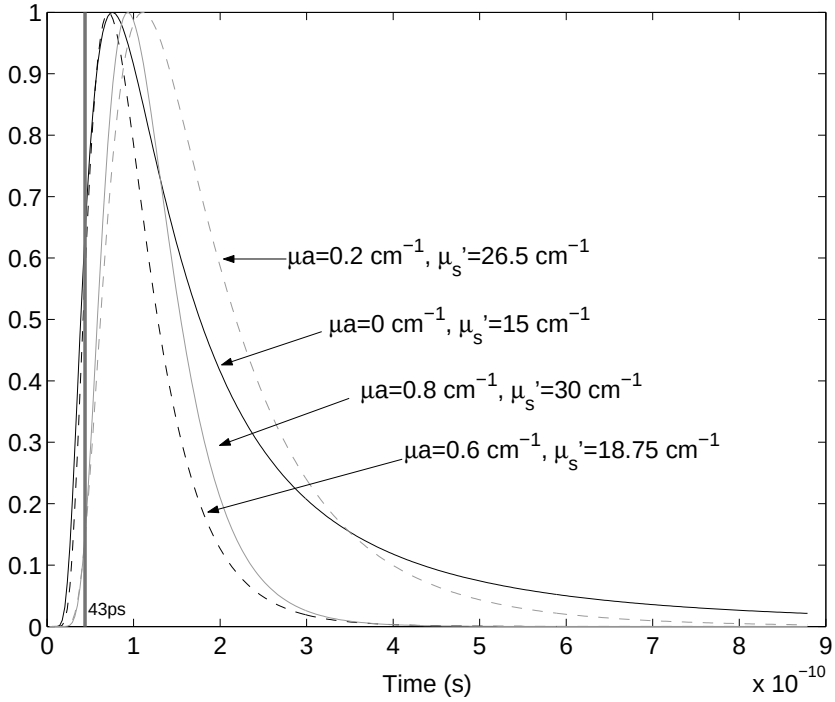


Figure 6: Measured signal and fitted signals for 3 wavelengths

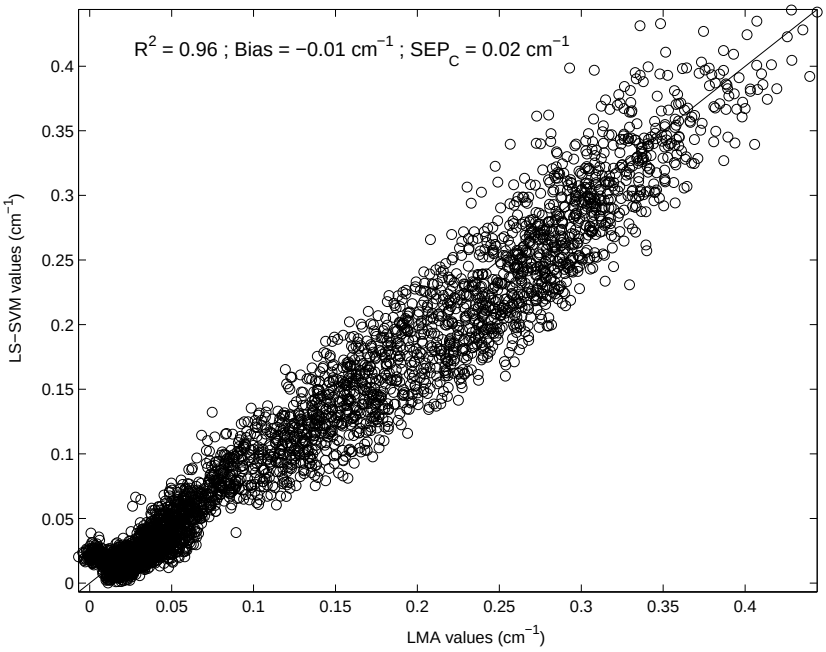


Figure 7: Prediction performances of μ_a prediction model

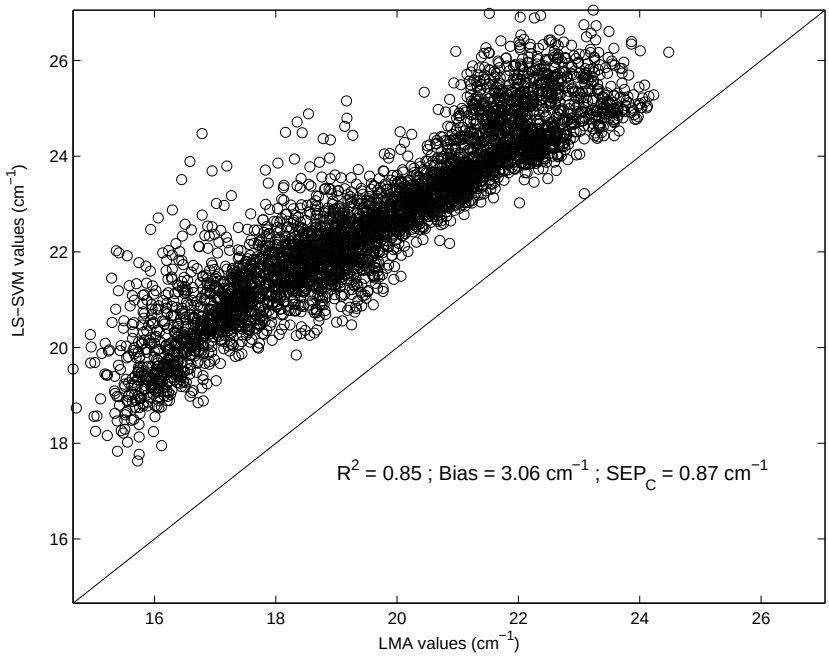


Figure 8: Prediction performances of μ_s' prediction model

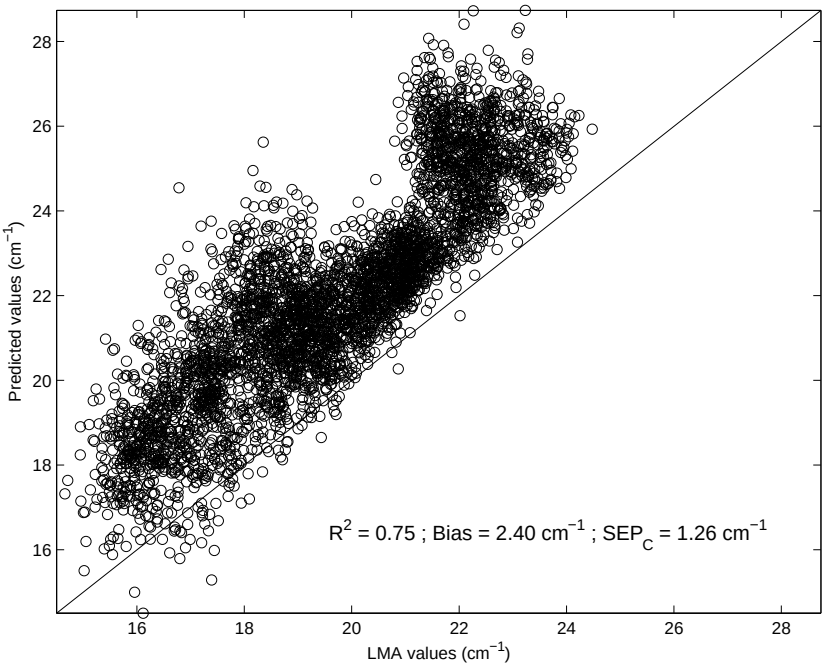


Figure 9: Prediction performances of μ_s' prediction model with convolution approach

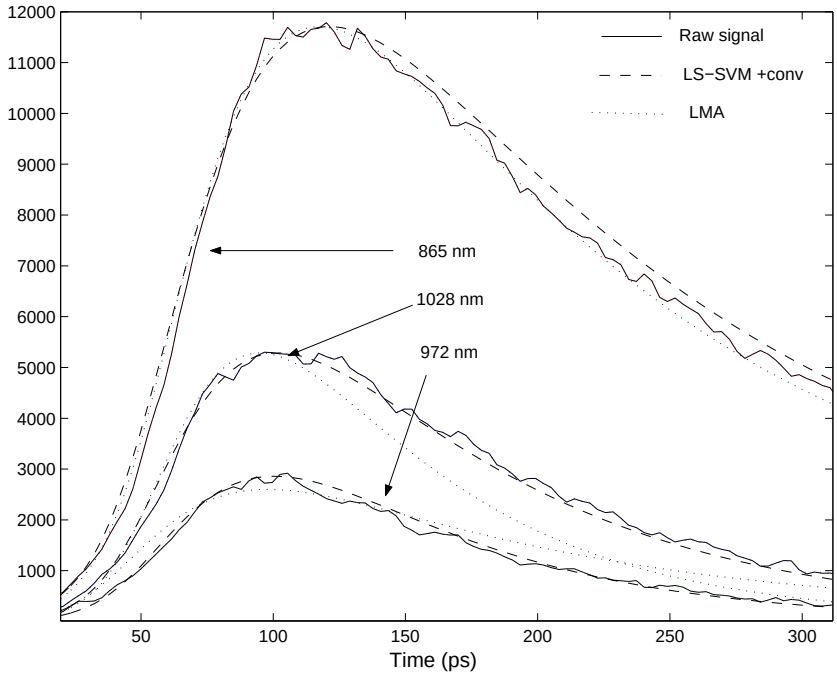


Figure 10: Measured signal and fitted signals for 3 wavelengths

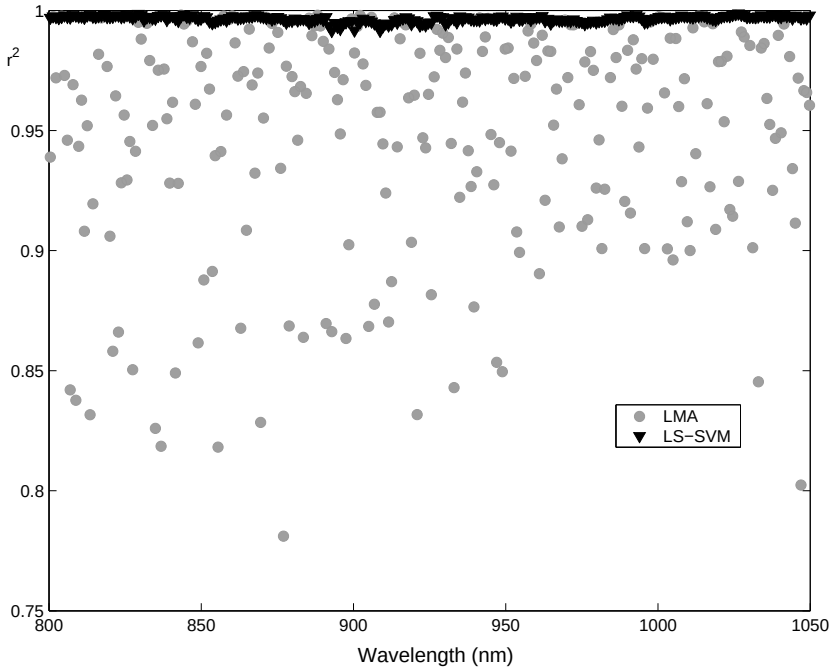


Figure 11: Determination coefficient between the raw temporal signal curve and one obtained using one of the two methods (LMA or LS-SVM) for each wavelength and for a given apple.