Combination of Models Obtained by Regression in the Wavelet Domain[1]

L.A. PINTO[2], Instituto Tecnológico de Aeronáutica, Divisão de Engenharia Eletrônica, 12228-900 São José dos Campos, SP, Brasil.
Instituto Federal do Espírito Santo, Coordenadoria de Engenharia de Controle e Automação, 29164-231 Serra, ES, Brasil.

R.K.H. GALVÃO[3], Instituto Tecnológico de Aeronáutica, Divisão de Engenharia Eletrônica, 12228-900 São José dos Campos, SP, Brasil.

Abstract. The wavelet transform is a useful tool to preprocess and compress datasets for linear regression modelling. However, the prediction performance of the resulting model depends on the choice of wavelet filter and number of decomposition levels, which may not be a straightforward task. This paper proposes an alternative approach, which consists of combining models obtained from different wavelet decompositions of the dataset. For this purpose, a method is developed to convert wavelet regression models back to the original domain. The proposed approach is illustrated in a case study involving the determination of density in gasoline samples by using infrared spectroscopy. The results are favourably compared to those obtained by using individual wavelet decompositions.

Keywords. Wavelet Transform, Combination of Models, Multiple Linear Regression.

1. Introdution

The wavelet transform (WT) has been used in a variety of signal processing applications, such as filtering [10], compression [24] and classification [8]. In particular, the WT has become a popular tool in a field of Analytical Chemistry known as Multivariate Calibration (MC) [7, 15]. The MC problem consists of building a mathematical model for estimation of physical or chemical properties of a sample from indirect measurements. Such measurements could be, for instance, emission, absorption or reflection intensities acquired over different wavelengths [20]. Applications of MC include determination of metals in steel alloys [16], analysis of composition of pharmaceutical formulae [5] and prediction of fuel properties [3], among many others.

---

[1]This work was supported by CNPq.

[2]luiz.pt@ifes.edu.br

[3]kawakami@ita.br

In the MC context, the WT is typically implemented by using filter bank algorithms [14]. Models are then built by applying regression techniques to the wavelet coefficients resulting from the bank filter decomposition [4, 6, 23]. In many applications, such a procedure led to improvements in the prediction performance of the resulting models [24]. However, the result usually depends on the choice of wavelet filter (Daubechies, Coiflet, or Symlet, for example [22] and number of decomposition levels. In fact, such choices are still an open problem in the MC literature. Usually, the authors refer to results previously obtained for similar datasets [7, 23] or choose one out of several combinations of wavelets filters and decomposition levels according to the performance of the resulting model [1, 15].

In view of these difficulties, the present paper proposes a new approach, which consists of combining models obtained from different wavelet decompositions. For this purpose, a method is developed to convert wavelet regression models back to the original domain. In this manner, the combination can be carried out by averaging the regression coefficients of the individual models. The proposed approach is illustrated in a study case involving the determination of density in gasoline samples by using infrared spectroscopy.

2.   The Wavelet Transform Implementation

The continuous wavelet transform (CWT) of a signal $x \in L^2(\mathbb{R})$ can be defined as

$$CWT(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(\lambda)\psi\left(\frac{\lambda - b}{a}\right) d\lambda,$$

where $\psi \in L^2(\mathbb{R})$ is the mother wavelet function and $a, b \in \mathbb{R}$ ($a \neq 0$) are the scale and translation parameters, respectively [22]. The magnitude of $CWT(a,b)$ is associated with the spectral content of $x(\lambda)$ around $\lambda = b$. Small scales correspond to high frequency values and vice-versa.

The discrete wavelet transform (DWT) can be obtained from the CWT by discretizing the scale and translation parameters as $a = a_0^j$ and $b = kb_0a_0^j$, where $j, k \in \mathbb{Z}$ and $a_0 > 1$, $b_0 > 0$ [22]. The DWT with $a_0 = 2$, $b_0 = 1$ (dyadic discretization) can be implemented in a computationally efficient manner by using a digital filter bank [14, 22]. The input sequence to such a filter corresponds to a sampled version of x, which will be denoted by $\mathbf{x}$.

Figure 1 presents the decomposition and reconstruction filter banks for calculation of the DWT and its inverse, respectively. In this example, a single decomposition level is employed. The basic structure of the decomposition filter bank consists of a pair of low-pass ($\mathbf{h}$) and high-pass ($\mathbf{g}$) filters, followed by a dyadic downsampling operation ($\downarrow 2$). The $\mathbf{g}$ filter can be associated to a mother wavelet. In the same manner, the $\mathbf{h}$ filter can be associated to a so-called "scaling function" [22]. The downsampled outputs of the low-pass and high-pass filters are termed approximation ($\mathbf{c}$) and detail ($\mathbf{d}$) coefficients, respectively. Sequences $\mathbf{c}$, $\mathbf{d}$ will henceforth be termed "wavelet coefficients". The filtering/downsampling operations can be reapplied to the approximation coefficients in order to obtain further approximations and details at lower resolution (i.e. larger scale) levels.

The **h** and **g** filters are typically of finite length, and therefore each approximation or detail coefficient corresponds to a section of the original signal. This spatial localization feature is one of the main advantages of the wavelet tranform over the Fourier transform [22].
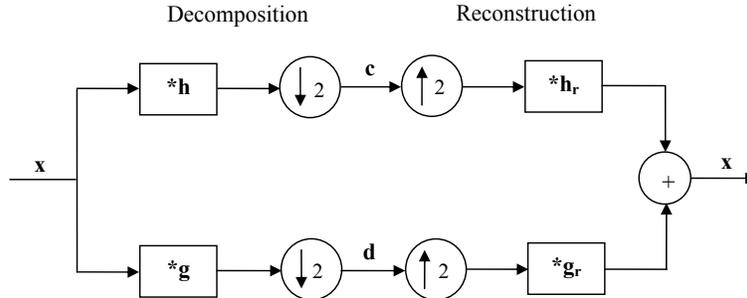

Figure 1: Filter bank implementation of the DWT (decomposition) and its inverse (reconstruction).

In the reconstruction filter bank, the approximation and detail coefficients undergo an operation in which zeros are inserted at the positions corresponding to the coefficients removed by the dyadic downsampling. The upsampled approximation and detail sequences are then filtered by lowpass ($\mathbf{h_r}$) and highpass ($\mathbf{g_r}$) reconstruction filters, respectively. In an orthogonal filter bank, these filters correspond to their decomposition counterparts in reverse order [6, 22].

The low-pass and high-pass filtering operations are carried out by convolution of the input signal with the filtering sequences (indicated by the $*$ symbol in Fig. 1). If the signal has finite length (which is usually the case), some boundary extension procedure must be employed [19, 22]. In the present work, the constant extension method (also termed smooth padding of order 0) is adopted [19]. Such a method leads to less pronounced border effects, as compared to periodic extension [6, 22] and zero padding [19]. However, the resulting transformation from the input signal to the wavelet coefficients is not orthogonal, even if an orthogonal filter bank is employed. Therefore, regression models in the wavelet domain cannot be converted back to the original domain by simple application of the reconstruction filter bank [22]. In the next section, a method is proposed to circumvent this problem.

3.   Conversion of Wavelet Regression Models to the Original Domain

For illustration, let $\mathbf{x} = [x_1 \ x_2 \ x_3 \ x_4 \ x_5]^T$ and $\mathbf{h} = [h_1 \ h_2 \ h_3 \ h_4]^T$ represent a signal of length five and a low-pass filter of length four, respectively. The convolution and downsampling operations result in four approximation coefficients $\mathbf{c} = [c_1 \ c_2 \ c_3 \ c_4]^T$. The mapping from $\mathbf{x}$ to $\mathbf{c}$ can be expressed in matrix form as

$\mathbf{c} = \mathbf{Hx}$, where

$$\mathbf{H} = \begin{bmatrix} (h_2 + h_3 + h_4) & h_1 & 0 & 0 & 0 \\ h_4 & h_3 & h_2 & h_1 & 0 \\ 0 & 0 & h_4 & h_3 & (h_1 + h_2) \\ 0 & 0 & 0 & 0 & (h_1 + h_2 + h_3 + h_4) \end{bmatrix}$$

assuming constant boundary extension. Similarly, the detail coefficients $\mathbf{d} = [d_1 \ d_2 \ d_3 \ d_4]^T$ can be obtained as $\mathbf{d} = \mathbf{Gx}$, where $\mathbf{G}$ is built from the high-pass filter weights $\mathbf{g} = [g_1 \ g_2 \ g_3 \ g_4]^T$.

A linear regression model in the wavelet domain can be written as

$$\hat{y} = \xi + (w_1 c_1 + w_2 c_2 + w_3 c_3 + w_4 c_4) + (z_1 d_1 + z_2 d_2 + z_3 d_3 + z_4 d_4) = \xi + \mathbf{w}^T \mathbf{c} + \mathbf{z}^T \mathbf{d} \quad (3.1)$$

where $\hat{y}$ is an estimate of the dependent variable y, ξ is an offset term and $\mathbf{w} = [w_1 \ w_2 \ w_3 \ w_4]^T$, $\mathbf{z} = [z_1 \ z_2 \ z_3 \ z_4]^T$ are the regression coefficients. In view of the relations $\mathbf{c} = \mathbf{Hx}$ and $\mathbf{d} = \mathbf{Gx}$, (3.1) can be rewritten as

$$\hat{y} = \xi + (\mathbf{w}^T \mathbf{H} + \mathbf{z}^T \mathbf{G})\mathbf{x}. \quad (3.2)$$

On the other hand, the regression model in the original domain would be of the form

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 + b_5 x_5 = b_0 + \mathbf{b}^T \mathbf{x}, \quad (3.3)$$

where $\mathbf{b} = [b_1 \ b_2 \ b_3 \ b_4 \ b_5]^T$. Therefore, from (3.2) and (3.3), it follows that

$$\xi + (\mathbf{w}^T \mathbf{H} + \mathbf{z}^T \mathbf{G})\mathbf{x} = b_0 + \mathbf{b}^T \mathbf{x}.$$

Since this identity must hold for any $\mathbf{x} \in \mathbb{R}^5$, one must have $b_0 = \xi$ and

$$\mathbf{b} = \mathbf{H}^T \mathbf{w} + \mathbf{G}^T \mathbf{z} \quad (3.4)$$

Equation (3.4) can also be written as $\mathbf{b} = \mathbf{b^h} + \mathbf{b^g}$, where $\mathbf{b^h} = \mathbf{H^T w}$, $\mathbf{b^g} = \mathbf{G^T z}$.

In the general case, for a signal $\mathbf{x}$ of odd length equal to K and a low-pass filter of length L, the number of approximation coefficients will be $M = (K + L - 1)/2$. The transformation from $\mathbf{w} = [w_1 \ w_2 \ \cdots \ w_M]^T$ to $\mathbf{b^h} = \begin{bmatrix} b_1^h \ b_2^h \ \cdots \ b_K^h \end{bmatrix}^T$ can be expressed as $\mathbf{b^h} = \mathbf{H^T w}$, where

$$\mathbf{H}^T = \begin{bmatrix} \sum_{i=2}^L h_i & \sum_{i=4}^L h_i & \cdots & h_L & 0 & \cdots & 0 & \cdots & 0 & 0 \\ h_1 & h_3 & \cdots & h_{L-1} & 0 & \cdots & 0 & \cdots & 0 & 0 \\ 0 & h_2 & \cdots & h_{L-2} & h_L & \cdots & 0 & \cdots & 0 & 0 \\ 0 & h_1 & \cdots & h_{L-3} & h_{L-1} & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & h_5 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & h_4 & \cdots & h_L & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & h_3 & \cdots & h_{L-1} & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & h_1 + h_2 & \cdots & \sum_{i=1}^{L-2} h_i & \sum_{i=1}^L h_i \end{bmatrix}$$

As can be seen, $\mathbf{b^h}$ can be obtained applying the reconstruction filter bank to $\mathbf{w}$ and correcting the endpoints b_1^h and b_K^h as

$$b_1^h = w_1 \sum_{i=2}^L h_i + w_2 \sum_{i=4}^L h_i + \cdots + w_{(L/2)-1} \sum_{i=L-2}^L h_i + w_{L/2} h_L \quad (3.5)$$

$$b_K^h = w_{M+1-(L/2)} \sum_{i=1}^{2} h_i + w_{M+2-(L/2)} \sum_{i=1}^{4} h_i + \cdots + w_{M-1} \sum_{i=1}^{L-2} h_i + w_M \sum_{i=1}^{L} h_i$$

Coefficients b_1^g, b_2^g, ..., b_K^g can be obtained from z_1, z_2, ..., z_M in exactly the same manner by using the high-pass filter weights g_1, g_2, ..., g_L.

If the signal length K is even, the number of approximation coefficients will be $M = (K + L - 2)/2$. The expression for b_1^h is still given by (3.5), but b_K^h will be calculated as

$$b_K^h = w_{M+1-(L/2)} h_1 + w_{M+2-(L/2)} \sum_{i=1}^{3} h_i + \cdots + w_{M-1} \sum_{i=1}^{L-3} h_i + w_M \sum_{i=1}^{L-1} h_i$$

The procedure described above can be applied to obtain the regression coefficients in the original domain from the regression coefficients associated to the first decomposition level of the filter bank. If the filter bank comprises more than one decomposition level, this procedure must be applied in a recursive manner to obtain the regression coefficients at the level j from the regression coefficients at level $(j + 1)$, where $j = 0$ corresponds to the original domain.

4.   Combination of Wavelet Regression Model

The proposed method consists of two phases [17]. In the first phase ("Generation of Models"), individual regression models are constructed for N different configurations (wavelet filter and number of decomposition levels) of the filter bank. In the present work, the Stepwise Regression algorithm [9] is adopted for this purpose. After conversion to the original domain, the nth model obtained can be written as

$$\hat{y}^{(n)} = b_0^{(n)} + \sum_{k=1}^{K} b_k^{(n)} x_k \tag{4.1}$$

For instance, if two different wavelet filters (Daubechies 4 and 5, for example) and one up to three decomposition levels are employed, then $N = 2 \times 3 = 6$.

In the second phase ("Combination of Models"), the individual models are averaged in order to obtain an ensemble model of the form

$$\hat{y}^{av} = b_0^{av} + \sum_{k=1}^{K} b_k^{av} x_k, \tag{4.2}$$

where $b_0^{av} = (1/N) \sum_{n=1}^{N} b_0^n$ and $b_k^{av} = (1/N) \sum_{n=1}^{N} b_k^{(n)}$, $k = 1, ..., K$.

In this work, three model combination strategies are compared. "Strategy 1" consists of combining models obtained by using the same wavelet filter, but with different decomposition levels. "Strategy 2" consists of combining models obtained with the same number of decomposition levels, but with different wavelet filters. Finally, "Strategy 3" consists of combining models obtained by varying both the wavelet filters and the decomposition levels.

5.   Application Example

The proposed method was applied to the problem of estimating the density of
gasoline samples by infrared spectroscopy. The dataset consisted of 104 samples
with absorbance spectra acquired at 6443 wavelengths. In order to eliminate
undesirable baseline fluctuations, first derivative spectra were calculated by using a
Savitzky-Golay filter [2] with a 2nd-order polynomial and an 11-point window. For
illustration, Figures 2a and 2b present the original and derivative spectra of one of
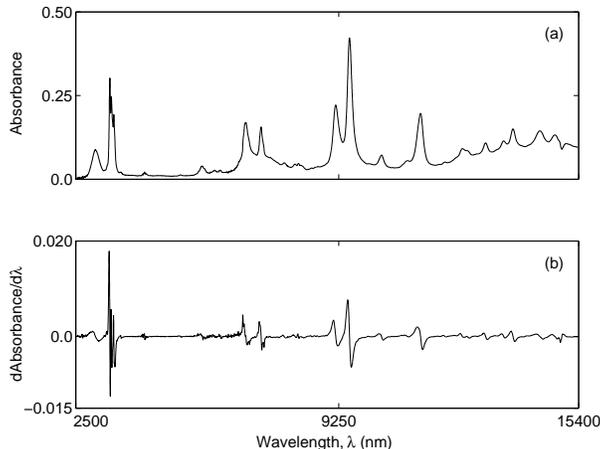the samples.

Figure 2: (a) Original and (b) derivative spectrum of a gasoline sample.

The Kennard-Stone algorithm [12, 13] was applied to select 73 calibration
samples, which were employed to build the regression models. The 31 remaining
samples formed a separate test set, which was used to compare the performance
of the resulting models. In order to reduce computational workload in the model-
building phase, a preliminary compression procedure was applied to the wavelet
coefficients. Compression was carried out by discarding the smallest wavelet
coefficients (in absolute value) while retaining 99% of the explained variance [18].

Stepwise Regression was employed with α-entry and α-exit values of 0.01 [9].
The individual models were constructed by employing 19 different filters from
the Daubechies (db2-db10), Symlet (sym4-sym8) and Coiflet (coif1-coif5) families
[1, 7, 15, 23]. In each case, the numbers of decomposition levels was varied from
one to a maximum of J. The value of J was set by imposing that at least one
wavelet coefficient should still refer to the original signal (without extension). For
the data set at hand, the decomposition levels corresponding to each wavelet filter
were the following: one up to eleven (db2), one up to ten (db3, coif1), one up to
nine (db4, db5, db6, coif2, sym4, sym5 and sym6), one up to eight (db7, db8, db9,
db10, coif3, coif4, sym7 and sym8) and one up to seven (coif5). Therefore, a total
of 165 individual models were obtained.

The resulting models were analyzed according to two metrics, namely: (*i*)

RMSEP (*root-mean-square error of prediction*) obtained in the test set and *(ii)* the 2-norm ($||b||_2$) of the regression vector. The RMSEP value is defined as

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

where $\hat{y}_i$, y_i denote the predicted and reference values of density (in g/cm^3) of the ith sample of the test set ($i = 1, \ldots, N_t$, with $N_t = 31$). The value of $||b||_2$ is an indicator of model sensitivity to noise in the **x** measurements [11, 21].

All calculations were carried out in the Matlab® 6.5 software (using functions from the Wavelet and Statistics Toolboxes), by using a computer with an Intel Core 2 Duo processor (1.86GHz, 3MB cache) and 3.5GB RAM.

6.   Results

Figure 3a presents the results obtained by using Strategy 1. Each circle corresponds to an ensemble model obtained for a given wavelet filter over different decomposition levels. In contrast, each square corresponds to the average result of individual models obtained for a given wavelet filter over different decomposition levels. The RMSEP values obtained by using Strategy 1 range from 0.8×10^{-3} to 1.0×10^{-3}, whereas the values for $||b||_2$ vary between 8.9 and 20.2. As can be seen, the proposed strategy leads to an improvement over the average results of the individual models, which range from 1.1×10^{-3} to 1.5×10^{-3} for RMSEP, and from 19.3 to 34.2 for $||b||_2$. It is worth noting that the results are further improved by using Strategy 3, as indicated by a star symbol in the plot. In this case, the RMSEP and $||b||_2$ values are 0.8×10^{-3} and 6.3 (that is, Strategy 3 matches the best RMSEP obtained by Strategy 1, with a smaller $||b||_2$ value).

The results obtained by using Strategy 2 are presented in Figure 3b. In this case, each circle corresponds to an ensemble model obtained for a given decomposition level over different wavelet filters. For comparison, each square corresponds to the average result of individual models obtained for a given decomposition level over different wavelet filters. Again, the ensemble models (RMSEP ranging from 0.7×10^{-3} to 1.1×10^{-3} and $||b||_2$ ranging from 6.6 to 22.5) can be seen to outperform the individual ones (RMSEP ranging from 1.1×10^{-3} to 1.5×10^{-3} and $||b||_2$ ranging from 16.9 to 40.2). It is worth noting that the best RMSEP for the ensemble models (0.7×10^{-3}) is smaller than the RMSEP value obtained with Strategy 3 (0.8×10^{-3}). However, $||b||_2$ is slightly smaller for Strategy 3.

In terms of computational workload, the three proposed strategies can be compared as follows. The construction of each individual model entails the wavelet decomposition of the derivative spectra, the preliminary compression applied to the wavelet coefficients, the Stepwise Regression procedure and the conversion of the model back to the original domain. On average, such operations were completed in 2.2 seconds. The overall time required to build an ensemble model in Strategy 1 ranged from 14 to 22 seconds, depending on the number of different decomposition levels employed in the combination. In Strategy 2, the time ranged
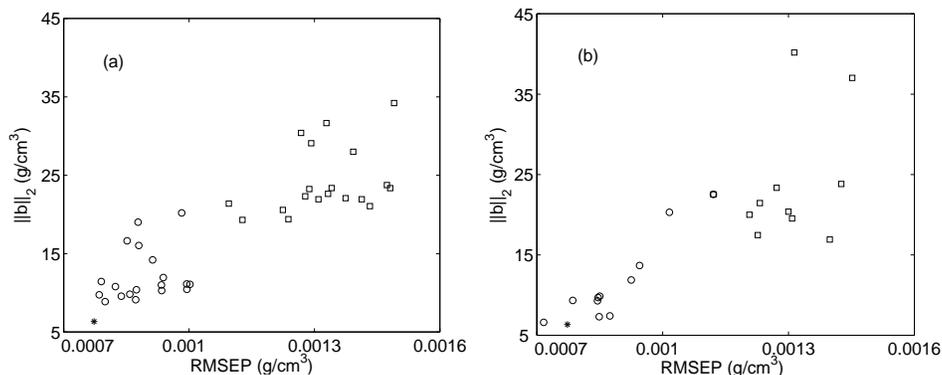
Figure 3: (a) Results of ensemble models obtained by using Strategy 1 (O) and Strategy 3 (*). Each □ symbol indicates the average result of individual models obtained for a given wavelet filter over different decomposition levels. (b) Results of ensemble models obtained by using Strategy 2 (O) and Strategy 3 (*). Each □ symbol indicates the average result of individual models obtained for a given decomposition level over different wavelet filters.

from 4 to 38 seconds, depending on the number of different wavelets employed in the combination. Strategy 3, in which 165 individual models are combined, required a total of 364 seconds.

As can be seen, the proposed model combination strategies are more time-consuming than the conventional approach of using a single wavelet filter and decomposition level. However, after an ensemble model has been constructed, the time required for subsequent prediction calculations (Equation 4.2) is the same as that of an individual model (Equation 4.1).

Finally, although Strategy 3 requires a larger computational workload, practical aspects may favour its adoption over Strategies 1 and 2. In fact, Strategy 1 still requires the analyst to choose a particular wavelet filter for the model-building process, whereas Strategy 2 requires the *a priori* choice of the number of resolution levels. In this sense, Strategy 3 is an "easier-to-use" approach, which provides results (RMSEP and $||b||_2$) similar to those of the best ensemble models obtained by Strategies 1 and 2 in the present study.

For illustration, Figure 4 presents the density predictions obtained by using Strategy 3. As can be seen, there is good agreement between predicted and reference values, both in the calibration and test data sets.

7.   Conclusions

This paper reported two contributions, namely a procedure to convert wavelet regression models back to the original domain, and a method for generating ensemble models by using filter banks with different configurations. The proposed method can be used as an alternative to the selection of a particular wavelet filter
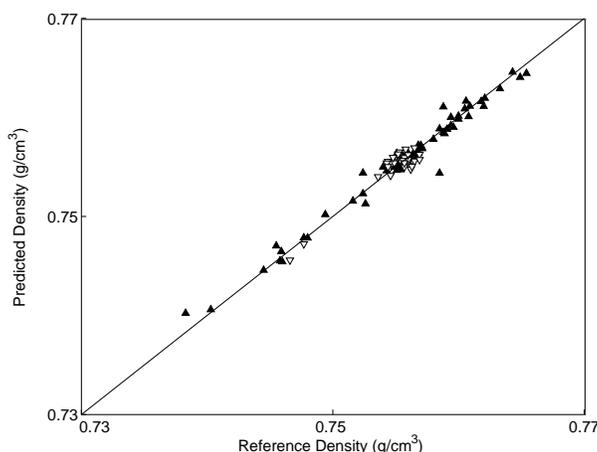
Figure 4: Prediction results for Strategy 3: calibration (▲) and test (▽) data sets.

and an appropriate number of decomposition levels, which is still a matter of much discussion in the literature.

For illustration, a case study involving the determination of density in gasoline by infrared spectroscopy was presented. In this application example, the ensemble models provided better results, in terms of prediction error and noise sensitivity, as compared to individual models obtained by the conventional approach.

Future works could be concerned with criteria to ascribe different weights to the models in the combination. Such weights could be related to the performance of the individual models as assessed, for instance, by cross-validation.

Acknowledgments

Resumo. A Transformada Wavelet é uma ferramenta útil para pré-processamento e compressão de dados em calibração multivariada. Contudo, a sua aplicação requer a escolha da wavelet e do número de níveis de decomposição a serem empregados, o que pode não ser uma tarefa simples. Este artigo propõe uma abordagem alternativa, que consiste em combinar modelos de regressão calibrados a partir de diferentes decomposições wavelet. Para isso, foi desenvolvido um método para converter ao domínio original modelos de regressão calibrados no domínio wavelet. Como ilustração, apresenta-se um estudo de caso envolvendo a determinação da densidade de gasolina por espectroscopia no infravermelho. Os resultados indicam que a abordagem proposta é uma alternativa vantajosa ao uso de uma única decomposição wavelet para construção do modelo.

References

[1] B.K. Alsberg, A.M. Woodward, M.K. Winson, J.J. Rowland, D.B. Kell, Variable selection in wavelet regression models, *Anal. Chim. Acta*, **368** (1998), 29–44.

[2] K.R. Beebe, R.J. Pell, B. Seasholtz, "Chemometrics - A Practical Guide", Wiley, New York, 1998.

[3] G. Bohacs, Z. Ovadi, A. Salco, Prediction of gasoline properties with near infrared spectroscopy, *J. Near Infrared Spectroscopy*, **6** (1998), 341–348.

[4] P.J. Brown, T. Fearn, M. Vannucci, Bayesian wavelet regression on curves with application to a spectroscopic calibration problem, *J. Am. Stat. Assoc.*, **96** (2001), 398–408.

[5] P. Chalus, S. Walter, M. Ulmschneider, Combined wavelet transform-artificial neural network use in tablet active content determination by near-infrared spectroscopy, *Anal. Chim. Acta,* **591** (2007), 219–224.

[6] C.J. Coelho, R.K.H. Galvão, M.C.U. Araújo, M.F. Pimentel, E.C. Silva, A solution to the wavelet transform optimization problem in multicomponent analysis, *Chemom. Intell. Lab. Syst.*, **66** (2003), 205–217.

[7] I.E. Díez, J.M. Saiz, C. Pizarro, OWAVEC: a combination of wavelet analysis and an orthogonalization algorithm as a pre-processing step in multivariate calibration, *Anal. Chim. Acta,* **515** (2004), 31–41.

[8] D. Donald, D. Coomans, Y. Everingham, D. Cozzolino, M. Gishen, T. Hancock, Adaptive wavelet modeling of a nested 3 factor experimental design in NIR chemometrics, *Chemom. Intell. Lab. Syst.*, **82** (2006), 122–129.

[9] N.R. Draper, H. Smith, "Applied Regression Analysis", Wiley, New York, 1998.

[10] R.K.H. Galvão, H.A. Dantas Filho, M.N. Martins, M.C.U. Araújo, C. Pasquini, Sub-optimal wavelet denoising of coaveraged spectra employing statistics from individual scans, *Anal. Chim. Acta,* **581** (2007), 159–167.

[11] J.H. Kalivas, Pareto calibration with built-in wavelength selection, *Anal. Chim. Acta,* **505** (2004), 9–14.

[12] K.R. Kanduc, J. Zupan, N. Majcen, Separation of data on the training and test set for modeling: a case study for modeling of five colours properties of a white pigment, *Chemom. Intell. Lab. Syst.*, **65** (2003), 221–229.

[13] R.W. Kennard, L.A. Stone, Computer aided design of experiment, *Technometrics*, **11** (1969), 137–148.

[14] S.G. Mallat, A theory for multiresolution signal decomposition: The wavelet representation, *IEEE Trans. Pattern Anal. Machine Intell.*, **11** (1989), 674–693.

[15] B.M. Nicolai, K.I. Theron, J. Lammertyn, Kernel PLS regression on wavelet transformed NIR spectra for prediction of sugar content of apple, *Chemom. Intell. Lab. Syst.*, **85** (2007), 243–252.

[16] M.F. Pimentel, B.B. Neto, M.C.U. Araújo, C. Pasquini, Simultaneous multielemental determination using a low-resolution inductively coupled plasma spectrometer/diode array detection system, *Spectrochimica Acta B.*, **52** (1997), 2151–2161.

[17] L.A. Pinto, R.K.H. Galvão, Combinação de modelos obtidos por regressão no domínio wavelet, *Anais do XXXII CNMAC*, **2** (2009), 761–767.

[18] M.J.C. Pontes, J. Cortez, R.K.H. Galvão, C. Pasquini, M.C.U. Araújo, R.M. Coelho, M.K. Chiba, M.F. Abreu, B.E. Madari, Classification of Brazilian soils by using LIBS and variable selection in the wavelet domain, *Anal. Chim. Acta*, **642** (2009), 12–18.

[19] R.N.F. Santos, "Calibração Multivariada para Análise Espectrofotométricas Empregando Pacotes Wavelet e Mínimos-Quadrados Parciais", Tese de Mestrado, ITA, São José dos Campos, SP, 2006.

[20] D.A. Skoog, F.J. Holler, T.A. Nieman, "Princípios de Análise Instrumental", Bookman, Porto Alegre, 2002.

[21] F. Stout, M.R. Baines, J.H. Kalivas, Impartial graphical comparison of multivariate calibration methods and the harmony/parsimony tradeoff, *J. Chemometrics,* **20** (2006), 464–475.

[22] G. Strang, T. Nguyen, "Wavelet and Filter Banks", Cambridge Press, Wellesley, 1996.

[23] J. Trygg, S. Wold, PLS regression on wavelet compressed NIR spectra, *Chemom. Intell. Lab. Syst.*, **42** (1998), 209–220.

[24] B. Walczak, D.L. Massart, Wavelet packet transform applied to a set of signals: A new approach to the best-basis selection, *Chemom. Intell. Lab. Syst.*, **38** (1997), 39–50.