A Linear Semi-infinite Programming Strategy for Constructing Optimal Wavelet Transforms in Multivariate Calibration Problems

Clarimar José Coelho,[†] Roberto K. H. Galvão,[‡] Mário César U. de Araújo,*[,||] Maria Fernanda Pimentel,[§] and Edvan Cirino da Silva[||]

Universidade Federal da Paraíba, Depto de Química, Caixa Postal 5093, 58051-970, João Pessoa, PB, Brazil, Universidade Católica de Goiás, Depto de Ciência da Computação, 74605-010, Goiânia, GO, Brazil, Instituto Tecnológico de Aeronáutica, Divisão de Engenharia Eletrônica, 12228-900, S. J. dos Campos, SP, Brazil, and Universidade Federal de Pernambuco, CTG, Depto de Engenharia Química, 50740-521, Recife, PE, Brazil

Received December 18, 2002

A novel strategy for the optimization of wavelet transforms with respect to the statistics of the data set in multivariate calibration problems is proposed. The optimization follows a linear semi-infinite programming formulation, which does not display local maxima problems and can be reproducibly solved with modest computational effort. After the optimization, a variable selection algorithm is employed to choose a subset of wavelet coefficients with minimal collinearity. The selection allows the building of a calibration model by direct multiple linear regression on the wavelet coefficients. In an illustrative application involving the simultaneous determination of Mn, Mo, Cr, Ni, and Fe in steel samples by ICP-AES, the proposed strategy yielded more accurate predictions than PCR, PLS, and nonoptimized wavelet regression.

INTRODUCTION

Notation. Matrices are represented by bold capital letters, vectors by bold lowercase letters, and scalars by italic characters. Elements of a sequence or vector are denoted by italic characters with a subscript index. The superscript T means transposed. $\mathbf{X}$ and $\mathbf{T}$ are the matrices $(M\times J)$ of calibration data in the original and transformed domains, respectively. A bar on top of a matrix denotes that some of its columns have been removed by variable selection. A vertical bar (|) represents the concatenation of two vectors or matrices. The hat symbol ($\wedge$) indicates a predicted value and $\mathscr{R}$ denotes the set of real numbers.

The problem of multivariate calibration in spectrometry consists of building a model that allows the determination of the composition of a sample from spectral measurements taken at several wavelengths.[1] Let y be the concentration of a certain analyte in a sample, which is to be determined from spectral measurements $x_0, x_1, ..., x_{J-1}$ taken at J wavelengths. A linear model for the estimation of y has the form

$$y = w_0 x_0 + w_1 x_1 + ... + w_{J-1} x_{J-1} + e = \mathbf{x}_{1\times J}\,\mathbf{w}_{J\times 1} + e \tag{1}$$

where e is a residual. To obtain the vector of model weights $\mathbf{w}$ from the data of M calibration samples, suffice it to write

$$\begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^M \end{bmatrix} = \begin{bmatrix} \mathbf{x}^1 \\ \mathbf{x}^2 \\ \vdots \\ \mathbf{x}^M \end{bmatrix} \mathbf{w} + \begin{bmatrix} e^1 \\ e^2 \\ \vdots \\ e^M \end{bmatrix} \tag{2}$$

where y^i, $\mathbf{x}^i$, and e^i are respectively the known analyte concentration, the vector of spectral measurements and the modeling residual for the ith sample. In matrix form, eq 2 can be written as

$$\mathbf{y}_{M\times 1} = \mathbf{X}_{M\times J}\,\mathbf{w}_{J\times 1} + \mathbf{e}_{M\times 1} \tag{3}$$

The estimate $\hat{\mathbf{w}}$ that minimizes $\mathbf{e}^T\mathbf{e}$ (sum of the squared residuals) is given by

$$\hat{\mathbf{w}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \tag{4}$$

provided $\mathbf{X}^T\mathbf{X}$ is nonsingular.

Notice that the precision of the model predictions depends on the signal-to-noise ratio of the spectrometer, which usually can only be improved by increasing the integration time during spectrum acquisition or by technological improvements in the equipment. However, the propagation of noise from the instrumental response to the model predictions can be reduced if the collinearity between the columns of $\mathbf{X}$ is made small.[1−4] This can be achieved by using a reduced subset of wavelengths instead of the entire frequency band monitored by the spectrometer.[3−7] However, since too much information may be discarded in this process, it may be more appropriate to perform the analysis in a different domain,[8] by transforming $\mathbf{X}$ as

$$\mathbf{T}_{M\times J} = \mathbf{X}_{M\times J}\,\mathbf{V}_{J\times J} \tag{5}$$

* Corresponding author phone/fax: +55 83 2167438/2167437; e-mail: laqa@quimica.ufpb.br.
[†] Universidade Católica de Goiás.
[‡] Instituto Tecnológico de Aeronáutica.
[§] Universidade Federal de Pernambuco.
[||] Universidade Federal da Paraíba.

10.1021/ci025657d CCC: $25.00
Published on Web 04/15/2003

where **V** is a matrix associated to a change of variables. If this transform has information compression features, some of the new variables may be discarded without significant loss of relevant information.

In this context, the most widely used chemometrical tools are Principal Component Regression (PCR) and Partial Least Squares (PLS), which project the data onto the directions of largest variance in the space.[1] However, it should be pointed out that PCR/PLS models may be difficult to interpret, because the latent variables on which the regression is performed lack physical meaning. Moreover, since the entire spectrum is involved in the construction of each principal component, it is not straightforward to separate the influence of nonrelevant spectral variables from those associated to the phenomenon of interest.

In this sense, recent works[8−16] have shown that the wavelet transform (WT) may be an interesting preprocessing tool in multivariate calibration, because in WT each transformed variable (wavelet coefficient) is related to a limited region of the spectrum, instead of the full spectrum. Thus, by using a convenient selection algorithm, it is possible to isolate wavelet coefficients that are related only to informative spectral regions.[8−10] The selected WT coefficients can then be used to build a model that will not be affected by the nonrelevant spectral variables of the data set.

One limitation of the proposed methods for the use of WT in multicomponent analysis[8−16] is that the wavelet must be chosen a priori and is not adapted to the experimental data set. To overcome this inconvenience, the present work proposes a novel strategy to construct a WT adapted to the statistics of the data set, based on the optimization of the filter bank implementation of WT. Specifically, the signal-adapted filter design technique described by Moulin and co-workers[17] is extended to optimize the compression ability of WT for the multivariate calibration problem. After the optimization phase, the best wavelet coefficients to include in the regression model are chosen by using the Successive Projections Algorithm (SPA), which is a selection technique for collinearity avoidance that has been recently developed by our research group.[3,4]

To illustrate the advantages of the proposed strategy, the wavelet optimization method is applied to a multivariate data set obtained by a low-resolution ICP-AES aiming at the simultaneous determination of Mn, Mo, Cr, Ni, and Fe in steel samples. It is worth noting that WT is particularly adequate in the analysis of signals with sharp peaks,[18] such as those found in plasma emission spectra. It could be argued that, in data sets acquired with modern ICP-AES instruments, collinearity is not as severe as in data obtained by other instrumental techniques, such as UV−VIS or NIR spectrometry. However, in the present application it will be shown that the use of low-resolution optics causes spectral overlapping and collinearity problems that are not negligible.

THEORETICAL BASIS

Filter Bank Implementation of the Wavelet Transform. The wavelet transform of a data vector **x** can be obtained in a computationally efficient manner by using a filter bank structure[19,20] such as the one depicted in Figure 1. In this figure, *H* and *G* are moving average low-pass and high-pass filters with weights $\{h_0, h_1, ..., h_{2N-1}\}$ and $\{g_0, g_1, ..., g_{2N-1}\}$,

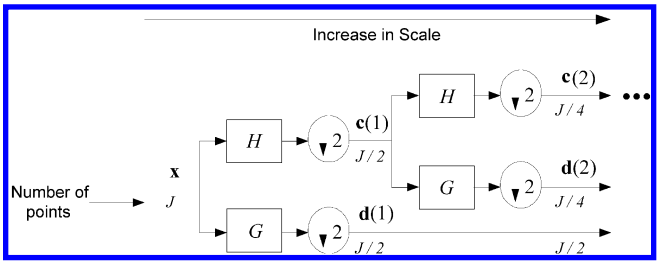

Figure 1. Filter bank implementation of the wavelet transform.

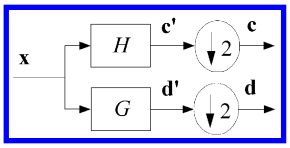

Figure 2. Low-pass/high-pass filter pair.

respectively, and (↓2) is the downsampling operation, which consists of removing every other point of a data sequence.

In the filter bank, the low-pass filtering result undergoes successive filtering iterations, up to a number of iterations *Nit* chosen by the analyst. The final result of the decomposition of data vector **x** is a vector $\mathbf{t} = [\mathbf{c}(Nit) \mid \mathbf{d}(Nit) \mid \mathbf{d}(Nit-1) \mid ... \mid \mathbf{d}(1)]$. Sequences **c**(k) and **d**(k) are termed respectively approximation and detail coefficients at the *k*th scale level. Coefficients in larger scales are associated to wider features in the data vector, whereas coefficients in smaller scales are associated to narrower features (such as sharp peaks).

In the optimization procedure, which will be proposed below, each low-pass/high-pass filter pair is optimized separately. Before describing the procedure, some considerations on the filtering process are needed. Figure 2 depicts the first low-pass/high-pass pair of the filter bank structure. In this figure, **c**′ and **d**′ represent, respectively, the approximation and detail coefficients before downsampling. In what follows, it will be considered that filtering is carried out by circular convolution.[21,22]

The use of circular convolution allows the filtering process to be represented in a simple matrix form as

$$\mathbf{c}' = \mathbf{h}\check{\mathbf{X}} \tag{6}$$

and

$$\mathbf{d}' = \mathbf{g}\check{\mathbf{X}} \tag{7}$$

where **h** and **g** are vectors formed from the low-pass and high-pass filter weights as

$$\mathbf{h} = [h_{2N-1}\, h_{2N-2} \ldots h_0] \tag{8}$$

and

$$\mathbf{g} = [g_{2N-1}\, g_{2N-2} \ldots g_0] \tag{9}$$

and $\check{\mathbf{X}}$ is a circulant matrix formed from the data vector **x** as

$$\check{\mathbf{X}} = \begin{bmatrix} x_0 & x_1 & \cdots & x_{J-2} & x_{J-1} \\ x_1 & x_2 & \cdots & x_{J-1} & x_0 \\ x_2 & x_3 & \cdots & x_0 & x_1 \\ \vdots & \vdots & & \vdots & \vdots \\ x_{2N-1} & x_{2N} & \cdots & x_{2N-3} & x_{2N-2} \end{bmatrix}_{2N \times J} \tag{10}$$

Notice that if the following conditions[19,20]

$$\sum_{n=0}^{2N-1-2l} h_n h_{n+2l} = \begin{cases} 1, & l = 0 \\ 0, & 0 < l < N \end{cases} \tag{11}$$

are satisfied and the high-pass and low-pass filter weights are related as

$$g_n = (-1)^{n+1} h_{2N-1-n} \tag{12}$$

then the decomposition structure is termed a quadrature-mirror filter (QMF) bank.[19,20] A QMF bank allows the data vector **x** to be perfectly reconstructed from the vector of wavelet coefficients **t**, meaning that there is no loss of information in the decomposition process.

The relation between the low-pass and high-pass filters expressed in eq 12 means that the transformation can be completely defined by the low-pass filter weights $\{h_0, h_1, ..., h_{2N-1}\}$. Since these 2*N* weights are subject to *N* restrictions (eq 11), there are *N* degrees of freedom that can be exploited to adjust WT to the statistics of the calibration data set. This adjustment can be carried out by defining an objective function related to the compression performance of the transform and using a search algorithm to optimize this function subject to the restrictions in eq 11. However, since the restrictions are nonlinear and may define a nonconvex search space, the optimization task is not trivial.

A possibility to circumvent this problem consists of parametrizing the low-pass filter weights by using a lattice structure with *N* angular parameters free of restrictions.[23,24] In this manner, search algorithms for unconstrained optimization could be employed. However, this approach is prone to problems of local maxima in the objective function, which makes the optimization outcome dependent on an initial choice of filter weights that must be made by the analyst. A parametrization technique that is not affected by local maxima problems is now proposed. It is based on a signal-adapted filter design technique[17] that allows the optimization to be restated as a linear semi-infinite programming problem.

Proposed WT Optimization Approach. The strategy proposed herein aims at concentrating the information in the approximation coefficients. This procedure is carried out by considering the power (energy divided by the number of coefficients) of the low-pass and high-pass filter outputs. Since the number of coefficients is *J*/2, due to the downsampling operation, the power of the approximation ($\mathbf{c}^i$) and of the detail ($\mathbf{d}^i$) coefficients for the *i*th calibration spectra are given by

$$P_c{}^i = \frac{\mathbf{c}^i \mathbf{c}^{iT}}{J/2} \tag{13}$$

$$P_d{}^i = \frac{\mathbf{d}^i \mathbf{d}^{iT}}{J/2} \tag{14}$$

If *M* calibration spectra are employed, the overall power of the approximations and details are

$$P_c = \sum_{i=1}^{M} P_c{}^i \tag{15}$$

$$P_d = \sum_{i=1}^{M} P_d{}^i \tag{16}$$

The relation between P_c and P_d can be expressed in an objective function F: $\mathscr{R}^{2N} \times \mathscr{R}^{2N} \rightarrow \mathscr{R}$ given by

$$F(\mathbf{h},\mathbf{g}) = \frac{0.5(P_c + P_d)}{\sqrt{P_c P_d}} \tag{17}$$

which is similar to the coding gain used to assess the compression performance of two-channel filter bank structures.[17,25]

If the conditions in eqs 11 and 12 are satisfied, WT is orthogonal, and thus it preserves power.[19] As a result, the power of the input signal $\mathbf{x}^i$ equals the sum of the power in both output channels, that is, $P_c{}^i + P_d{}^i$. Hence, for a given calibration set, the sum $P_c + P_d$ is constant and thus maximizing F is equivalent to maximizing P_c (see Appendix). It should be pointed out that the utility of maximizing P_c lies in the fact that the signal-to-noise ratio is usually larger in the low-pass filter output. Thus, maximizing P_c further improves the filtering performance. Since P_c only depends on the low-pass filter weights **h**, the problem can be restated as the maximization of an objective function ϵ: $\mathscr{R}^{2N} \rightarrow \mathscr{R}$ given by

$$\epsilon(\mathbf{h}) = P_c = \frac{1}{J/2} \sum_{i=1}^{M} \mathbf{c}^i \mathbf{c}^{iT} \tag{18}$$

Moreover, if the power of the detail coefficients before and after downsampling is approximately the same, function ϵ can be rewritten as

$$\epsilon(\mathbf{h}) = \frac{1}{J} \sum_{i=1}^{M} \mathbf{c}'^i \mathbf{c}'^{iT} \tag{19}$$

where $\mathbf{c}'^i$ is the vector of detail coefficients for the *i*th calibration signal, before the downsampling operation. From eq 6, it follows that $\mathbf{c}'^i = \mathbf{h}\check{\mathbf{X}}^i$, where $\check{\mathbf{X}}^i$ is the circulant matrix formed from $\mathbf{x}^i$. Thus, eq 19 becomes

$$\epsilon(\mathbf{h}) = \frac{1}{J} \sum_{i=1}^{M} \mathbf{h}\check{\mathbf{X}}^i \check{\mathbf{X}}^{iT} \mathbf{h}^T \tag{20}$$

or

$$\epsilon(\mathbf{h}) = \mathbf{h}\mathbf{R}\mathbf{h}^T \tag{21}$$

where

$$\mathbf{R} = \sum_{i=1}^{M} \left(\frac{1}{J} \check{\mathbf{X}}^i \check{\mathbf{X}}^{iT} \right) \tag{22}$$

Since $\check{\mathbf{X}}^i \check{\mathbf{X}}^{iT}$ is Toeplitz (a square matrix in which each diagonal parallel to the main diagonal is composed of equal elements[26]) for any $\mathbf{x}^i$, then $\mathbf{R}_{2N\times 2N}$ is also a Toeplitz matrix. Thus, the constraints in eq 11 allow the objective function to be rewritten, with a slight abuse of notation, in the following linear form[17]

$$\epsilon(\mathbf{a}) = \frac{r_0}{2} + \sum_{n=0}^{N-1} a_n r_{2n+1} \tag{23}$$

where $\{r_0, r_1, ..., r_{2N-1}\}$ are the elements of the first row of

R and vector $\mathbf{a} = [a_0\ a_1\ ...\ a_{N-1}]$ contains the coefficients of the product filter $P(z)$ defined as

$$P(z) = H(z)H(z^{-1}) = 1 + \sum_{n=0}^{N-1} a_n(z^{2n-1} + z^{2n+1}) \quad (24)$$

where

$$H(z) = \sum_{n=0}^{2N-1} h_n z^{-n} \quad (25)$$

is the discrete transfer function[21,22] of the low-pass filter.

Given **a**, the transfer function $H(z)$ of the desired filter can be recovered from $P(z)$ by a factorization procedure.[19] This factorization is always possible provided that the frequency response of the product filter $Q(f)$ given by

$$Q(f) = P(\exp(j2\pi f)),\ j = \sqrt{-1} \quad (26)$$

is non-negative at all frequencies f, that is, $Q(f) \geq 0$, $\forall f \in \mathscr{R}$. Substituting eq 24 in 26, $Q(f)$ can be written as

$$Q(f) = 1 + 2\sum_{n=0}^{N-1} a_n \cos[2\pi f(2n+1)] \quad (27)$$

Since $Q(f)$ is periodic with period 1 and $Q(f) + Q(f + 0.5) = 2$, $\forall f \in \mathscr{R}$, it is sufficient to consider the restriction $Q(f) \geq 0$ in the interval $0 \leq f \leq 0.5$, that is

$$1 + 2\sum_{n=0}^{N-1} a_n \cos[2\pi f(2n+1)] \geq 0,\ 0 \leq f \leq 0.5 \quad (28)$$

Maximizing $\epsilon(\mathbf{a})$ defined in eq 23 with respect to **a** subject to the inequality restrictions in eq 28 is a linear semi-infinite programming problem,[27] because there is a finite number of variables $(a_0\ a_1\ ...\ a_{N-1})$ and infinitely many restrictions. This problem can be solved by discretizing the frequency interval [0, 0.5] to generate a finite number of restrictions and then applying standard linear programming techniques.[28] The solution $\tilde{\mathbf{a}}$ to this approximated problem can then be used to generate a feasible solution $\mathbf{a_f}$ to the original problem as[17]

$$\mathbf{a_f} = \frac{\tilde{\mathbf{a}}}{1 - \delta} \quad (29)$$

where $\delta \leq 0$ is the minimum of $Q(f)$ in the interval $0 \leq f \leq 0.5$ when $\tilde{\mathbf{a}}$ is used instead of **a** in eq 27.

EXPERIMENTAL SECTION

The WT optimization technique was applied to steel spectra in the range 245−326 nm obtained with the low-resolution inductively coupled plasma-atomic emission spectrometer and methodology described elsewhere.[29,30] The XVERT algorithm[31] was used to design $M = 17$ calibration mixtures with element concentrations in the following ranges: Mn: 0.4−2.0%; Mo: 0.6−4.0%; Cr: 8−28%; Ni: 4−28%; Fe: 40−88%. Model assessment was based on a test set consisting of 10 synthetic mixtures (simulating typical steel samples) plus nine reference steel samples. A sample spectrum is presented in Figure 3. Notice that sharp spectral peaks are alternated with regions with low signal-to-noise ratio, which justifies the use of WT.

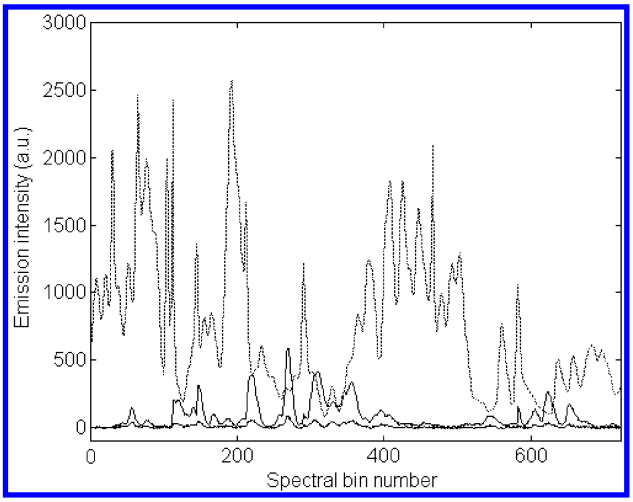

Figure 3. Spectra of 10 and 1.5 mg L^{-1} molybdenum standard solutions (solid lines) and of a sample (dashed line) containing 3.0, 6.0, 50.0, 40.0, and 150.0 mg L^{-1} of Mn, Mo, Cr, Ni, and Fe, respectively. The horizontal axis represents the wavelength index.

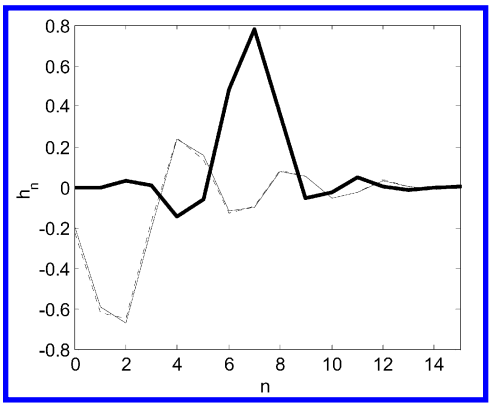

Figure 4. Optimized low-pass filter weights at the first (solid thin line) and second (dashed line) decomposition level, compared to the Symlet 8 low-pass filter weights (thick line).

The difficulties associated to the use of low-resolution optics in this application are also illustrated in Figure 3. As can be seen, the spectrum of Mo is strongly overlapped by the emission of the other elements. Moreover, the emission bands are spread across several wavelengths, which generates collinearity between the spectral variables.

RESULTS AND DISCUSSION

The spectra were decomposed with a two-level filter bank. The optimization was carried out on the first pair of filters, and then the second pair was optimized with respect to the approximation coefficients yielded by the first low-pass filter. In both decomposition levels, filters of length $2N = 16$ were used, as in the work of Alsberg and co-workers.[9,10] The frequency interval [0, 0.5] was partitioned in 1000 equally spaced points, and the well-known Simplex method[28] was used to solve the resulting linear programming problem.

Figure 4 presents the low-pass filter weights obtained at the two decomposition levels. The weights associated to the Symlet 8 wavelet, adopted by Alsberg and co-workers[9,10] are also shown for comparison. It is interesting to note that the optimized filter weights at the first and second level are

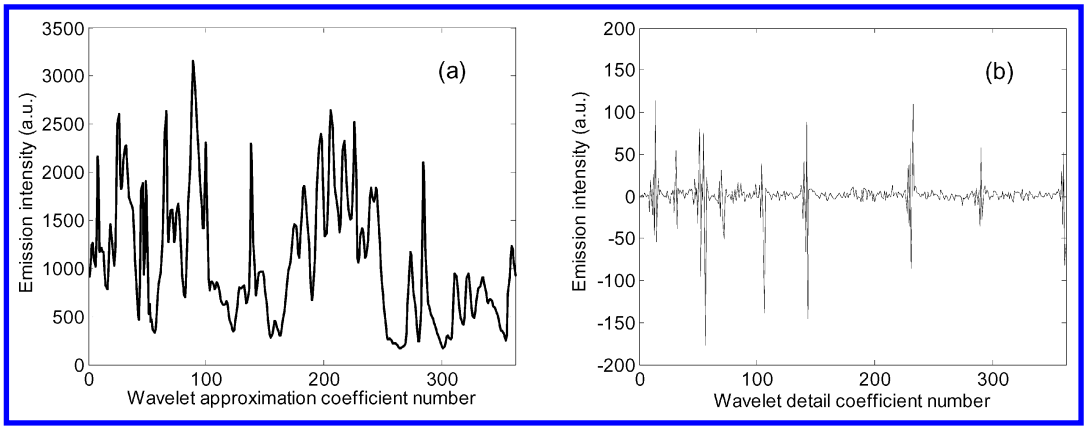

Figure 5. (a) Approximation coefficients and (b) detail coefficients obtained at the first decomposition level with the optimized WT applied to a sample spectrum.

similar. This finding can be explained by noting that the approximation coefficients after the first decomposition level are similar in shape to the original spectrum, as can be seen by comparing Figures 3 and 5a. The small differences are due to noise and very narrow peaks, which are extracted to the detail coefficients shown in Figure 5b. Furthermore, it should be noticed that the optimized filter weights are considerably different from the Symlet 8 weights, which suggests that the Symlet 8 wavelet might not be the best choice for this particular data set. In fact, by evaluating the objective function F in eq 17 for the first decomposition level, one obtains $F = 12.4$ for Symlet 8 and $F = 13.5$ for the optimized WT, representing a 9% improvement in filtering performance.

After optimization, the Successive Projections Algorithm[3,4] was used to choose the wavelet coefficients to include in the model. To speed up the selection, the coefficients that jointly explained less than 1% of the total data variance were initially removed. The use of SPA allowed the model to be built by multiple linear regression on the selected wavelet coefficients. The resulting model was compared to one obtained with a Symlet 8 decomposition,[9,10] by using the same number of levels (two) and the same variable selection algorithm. Comparisons were also made with Multiple Linear Regression (MLR) on the original spectral variables (with QR decomposition[2] to deal with ill-conditioning), PCR and PLS. For PCR and PLS, the number of factors was chosen by external validation and full cross-validation procedures in the UNSCRAMBLER Chemometrics software (CAMO A/S), version 6.1.

Table 1 presents the results in terms of the root-mean-square relative error of prediction (*RMSREP*) defined for each analyte as

$$RMSREP = \sqrt{\frac{1}{M_{test}}\sum_{i=1}^{M_{test}}\left[\frac{y^i - \hat{y}^i}{y^i}\right]^2} \qquad (30)$$

where y^i and $\hat{y}^i$ are respectively the true and predicted values of the analyte concentration in the ith test sample and M_{test} is the number of test samples.

As can be seen, regression in the original spectral domain leads to a model with poor prediction ability, due to the collinearity between the spectral variables, which was not removed satisfactorily by the QR algorithm. The result is considerably improved by the use of PCR and PLS, but the best models are obtained with WT. When the two wavelet models are compared, the advantage of using a WT optimized with respect to the statistics of the data set becomes apparent. In fact, the optimization leads to a decrease of 25% in the average *RMSREP*, apart from improving the general parsimony of the wavelet models.

Table 1. *RMSREP* (%) for the Test Set[a]

	analyte					
method	Mn	Mo	Cr	Ni	Fe	average
MLR-QR	10.7 (723)	9.2 (723)	5.3 (723)	5.1 (723)	2.0 (723)	6.5
PCR	2.2 (5)	4.6 (7)	1.0 (5)	2.0 (7)	1.1 (4)	2.2
PLS	2.6 (5)	4.0 (6)	1.0 (5)	2.1 (6)	1.0 (3)	2.1
Symlet 8 wavelet	1.6 (7)	3.2 (14)	0.9 (5)	1.6 (10)	0.6 (12)	1.6
Optimized WT	1.2 (7)	2.5 (8)	0.9 (17)	1.0 (8)	0.6 (7)	1.2

[a] The number of variables used in each model (spectral variables for MLR-QR, factors for PCR/PLS, and coefficients for the wavelet models) is shown inside parentheses.

CONCLUSIONS

This work proposed a novel strategy for optimizing a wavelet transform with respect to the statistics of the data set in multivariate calibration problems. The proposed method has the advantage of not being affected by local maxima problems. In this manner, it does not require the analyst to provide a convenient starting point for the optimization algorithm. Moreover, stochastic operations (such as those employed in genetic algorithms[5] or simulated annealing[32]) do not need to be employed in the search for the maximum of the objective function, thus ensuring that the optimization result is reproducible. The numerical processing effort involved is also modest, since most workload is carried out by the Simplex method for linear programming, which is a computationally efficient algorithm. Furthermore, the application of the resulting wavelet model to the analysis of new samples is straightforward. In fact, once the high-pass and low-pass weights have been obtained, the filter bank decomposition process is similar to that of a nonoptimized WT.

In the application presented for illustration, the wavelet optimization procedure led to a model with better prediction

ability than models obtained by MLR, PCR, PLS, and standard (nonoptimized) wavelet models. These results demonstrate that the proposed strategy is indeed an advantageous alternative to the use of nonoptimized WT in multivariate calibration problems. Future works could exploit the potentiality of the strategy in other chemometrics applications, such as pattern recognition and classification.

ACKNOWLEDGMENT

The authors thank FINEP-CTPETRO (Proc. 0652/00) science funding program, FAPESP (Grant 00/09390-6), CNPq (PRONEX Grant 015/98), and CAPES for partial financial support. The authors also wish to thank Dr. Sillas Hadjiloucas (Cybernetics Department, Reading University, U.K.) for his valuable comments. The research fellowships granted by the Brazilian agencies CNPq and CAPES are also gratefully acknowledged.

APPENDIX

Let A be the constant sum of approximation and detail power $P_c + P_d$ for a given calibration data set. Then the objective function in eq 17 can be rewritten as

$$F = \frac{0.5(P_c + P_d)}{\sqrt{P_c P_d}} = \frac{0.5A}{\sqrt{P_c(A - P_c)}} \tag{31}$$

or

$$F^2 = \frac{0.25A^2}{AP_c - P_c^2} \tag{32}$$

It can be seen that F reaches a minimum of 1 when $P_c = A/2$, that is, when the power is equally divided between the approximation and detail coefficients. As P_c increases from $A/2$ to A, F increases and tends to $+\infty$ when P_c tends to A, that is, when all power is contained in the approximation coefficients.

REFERENCES AND NOTES

(1) Martens, H.; Naes, T. *Multivariate Calibration*; John Wiley: New York, 1989.
(2) Lawson, C. L.; Hanson, R. J. *Solving Least-Squares Problems*; Prentice Hall: Englewood Cliffs, 1974.
(3) Galvao, R. K. H.; Pimentel, M. F.; Araujo, M. C. U.; Yoneyama, T.; Visani, V. Aspects of the successive projections algorithm for variable selection in multivariate calibration applied to plasma emission spectrometry. *Anal. Chim. Acta* **2001**, *443*, 107−115.
(4) Araujo, M. C. U.; Saldanha, T. C. B.; Galvão, R. K. H.; Yoneyama, T.; Chame, H. C.; Visani, V. The successive projections algorithm for variable selection in spectroscopic multicomponent analysis. *Chemomet. Intell. Lab. Syst.* **2001**, *57*, 65−73.
(5) Jouan-Rimbaud, D.; Massart, D. L.; Leardi, R.; Noord, O. E. Genetic algorithms as a tool for wavelength selection in multivariate calibration. *Anal. Chem.* **1995**, *67*, 4295−4301.
(6) Sasaki, K.; Kawata, S.; Minami, S. Optimal wavelength selection for quantitative analysis. *Appl. Spectrosc.* **1986**, *40*, 185−190.
(7) Xu, L.; Schechter, I. Wavelength selection for simultaneous spectroscopy analysis. Experimental and theoretical study. *Anal. Chem.* **1996**, *68*, 2392−2400.
(8) Brown, P. J.; Fearn, T.; Vannucci, M. Bayesian wavelet regression on curves with application to a spectroscopic calibration problem. *J. Am. Stat. Assoc.* **2001**, *96*, 398−408.
(9) Alsberg, B. K.; Woodward, A. M.; Kell, D. B. An introduction to wavelet transforms for chemometricians: A time-frequency approach. *Chemom. Intell. Lab. Syst.* **1997**, *37*, 215−239.
(10) Alsberg, B. K.; Woodward, A. M.; Winson, M. K.; Rowland, J. J.; Kell, D. B. Variable selection in wavelet regression models. *Anal. Chim. Acta* **1998**, *368*, 29−44.
(11) Depczynski, U.; Jetter, K.; Molt, K.; Niemoller, A. Quantitative analysis of near-infrared spectra by wavelet coefficient regression using a genetic algorithm. *Chemomet. Intell. Lab. Syst.* **1999**, *47*, 179−187.
(12) Jetter, K.; Depczynski, U.; Molt, K.; Niemoller, A. Principles and applications of wavelet transformation of chemometrics. *Anal. Chim. Acta* **2000**, *420*, 169−180.
(13) Walczak, B. *Wavelets in Chemistry*; Elsevier Science: New York, 2000.
(14) Eriksson, L.; Trygg, J.; Johansson, E.; Bro, R.; Wold, S. Orthogonal signal correction, wavelet analysis, and multivariate calibration of complicated process fluorescence data. *Anal. Chim. Acta* **2000**, *420*, 181−195.
(15) Trygg, J.; Wold, S. PLS regression on wavelets compressed NIR spectra. *Chemomet. Intell. Lab. Syst.* **1998**, *42*, 209−220.
(16) Vogt, F.; Tacke, M. Fast principal component analysis of large data sets. *Chemom. Intell. Lab. Syst.* **2001**, *59*, 1−18.
(17) Moulin, P.; Anitescu, M.; Kortanek, K. O.; Potra, F. A. The role of linear semi-infinite programming in signal-adapted qmf bank design. *IEEE Trans. Signal Proc.* **1997**, *45*, 2160−2174.
(18) Pen, U. L. Application of wavelets to filtering of noisy data. *Philos. Trans. R. Soc. London A − Math. Phys. Eng. Sci.* **1999**, *357*, 2561−2571.
(19) Strang G.; Nguyen, T. *Wavelets and Filter Banks*; Wellesley-Cambridge Press: Wellesley, 1996.
(20) Daubechies, I. *Ten Lectures on Wavelets*; Society for Industrial and Applied Mathematics: Philadelphia, 1992.
(21) Proakis, J. G.; Manolakis, D. G. *Digital Signal Processing: Principles, Algorithms and Applications*, 3rd ed.; Prentice Hall: Englewood Cliffs, 1995.
(22) Oppenheim, A. V.; Schafer, R. W. *Discrete-Time Signal Processing*; Prentice Hall: Englewood Cliffs, 1989.
(23) Sherlock, B. G.; Monro, D. M. On the space of orthonormal wavelets. *IEEE Trans. Signal Proc.* **1998**, *46*, 1716−1720.
(24) Vaidyanathan, P. P. *Multirate Systems and Filter Banks*; Prentice Hall: Englewood Cliffs, 1992.
(25) Unser, M. On the optimality of ideal filters for pyramid and wavelet signal approximation. *IEEE Trans. Signal Proc.* **1993**, *41*, 3591−3596.
(26) Barnabei, M.; Guerrini, C.; Montefusco, L. B. Some algebraic aspects of signal processing. *Linear Algebra Applications* **1998**, *284*, 3−17.
(27) Hettich, R.; Kortanek, K. O. Semi-infinite programming − theory, methods, and applications. *SIAM Rev.* **1993**, *35*, 380−429.
(28) Chvatal, V. *Linear Programming*; Freeman: New York, 1983.
(29) Pimentel, M. F.; Neto, B. B.; Araújo, M. C. U.; Pasquini, C. Simultaneous multielemental determination using a low-resolution inductively coupled plasma spectrometer/diode array detection system. *Spectrochim. Acta B* **1997**, *52*, 2151−2161.
(30) Pimentel, M. F.; Araujo, M. C. U.; Neto, B. B.; Pasquini, C. Conversion of a sequential inductively coupled plasma emission spectrometer into a multichannel simultaneous system using a photodiode array detector. *J. Autom. Chem.* **1998**, *20*, 69−75.
(31) Cornell, J. How to run mixture experiments for product quality. *The ASQC Basic References in Quality Control: Statistical Techniques*; American Society for Quality Control: Milwaukee, 1990.
(32) Kalivas, J. H.; Roberts, N.; Sutter, J. M. Global optimization by simulated annealing with wavelength selection for ultraviolet−visible spectrophotometry. *Anal. Chem.* **1989**, *61*, 2024−2030.

CI025657D