

# Determination of water in ferrous lactate by near infrared reflectance spectroscopy with a fibre-optic probe

M. Blanco \*, J. Coello, H. Iturriaga, S. MasPOCH, E. Rovira

*Departamento de Química, Unidad de Química Analítica, Universidad Autónoma de Barcelona, E-08193 Bellaterra, Barcelona, Spain*

Received 1 August 1996; received in revised form 20 January 1997

---

## Abstract

Near infrared diffuse reflectance spectroscopy with a fibre-optic probe was used to determine the water content in ferrous lactate dihydrate. Spectra were recorded by immersing the probe in a beaker containing the ferrous lactate sample. Spectral data were processed by using two different multivariate calibration procedures, viz. stepwise multiple linear regression (SMLR) and partial least-squares regression (PLSR). The results provided by the two calibration procedures were similar and departed by less than 1.5% from the values obtained by Karl Fischer titration. © 1997 Elsevier Science B.V.

*Keywords:* Fiber-optics; Multivariate calibration; Near infrared reflectance spectroscopy; Pharmaceutical preparations; Water determination

---

## 1. Introduction

The presence of crystallization or hydration water in solids markedly alters some properties of special interest to the pharmaceutical industry such as the rate of chemical degradation, crystal dimensions, solubility, compaction power, etc. Hence, not surprisingly, the determination of the water content in both raw materials and end products is one of the most commonplace analyses and as such calls for expeditious, reliable analytical methods.

The Pharmacopeia [1] recommends thermogravimetry and the Karl Fischer method for determining the water content. The former is a very accurate but also very sluggish technique that can easily take several hours to complete. Titration with the Karl Fischer reagent is much faster, so it is the more usual procedure. However, it is subject to some interferences [2], so it cannot be applied to any type of sample. Also, like any volumetric method, it calls for highly precise weighing, so it is scarcely appropriate for plant control measurements.

The determination of water in pharmaceutical preparations was one of the earliest applications of near infrared spectroscopy (NIRS) [3]. The shortcomings of the initial procedure, which en-

---

\* Corresponding author. Tel.: +34 3 5811367; fax: +34 3 5812477; e-mail: IQAN1@CC.UAB.ES.

tailed dissolving the sample in a solvent, and those of the instruments available at the time, designed to work mainly in transmittance mode, significantly restricted their use. The subsequent development of near infrared diffuse reflectance spectroscopy (NIRRS) [4], directly applicable to solid samples with no pretreatment, enhanced the analytical interest in this technique, especially on account of its expeditiousness and applicability to any type of sample, which allows its ready adaptation to control analyses [5,6].

A NIRRS signal is a complex function that depends on both physical (grain size, crystal structure) and chemical properties of the sample. At any wavelength, there is a linear relation similar to that of the Lambert–Beer law between the reflectance of a sample and its concentration:

$$\log \frac{R'}{R} = \frac{aC}{s} \quad (1)$$

where  $R'$  denotes the reflectance of a non-absorbing standard,  $R$  that of the sample,  $a$  the absorptivity,  $C$  the concentration and  $s$  the scattering coefficient. The logarithm of  $R'$  is constant and can be neglected, so, for a single component, Eq. (1) can be rewritten as

$$C = K + \frac{s}{a} \log \frac{1}{R} \quad (2)$$

By analogy with the Lambert–Beer law,  $\log(1/R)$  is known as the ‘apparent absorbance’.

One of the features of NIRRS is its ability to make measurements by using a fibre-optic module in order to considerably increase the throughput. The chief shortcoming of using a fibre-optic probe is that it increases spectral noise; as a result, most reported applications involve qualitative analysis, particularly for the identification of raw materials [7].

However, comparative studies between quantitative analyses were recently carried out by using the conventional cuvette module and a fibre-optic module; the results were quite comparable [8], provided spectral data at wavelengths above 2200 nm—where fibre optics exhibits the highest spectral noise—were discarded.

Ferrous lactate is used by the pharmaceutical industry as a source of therapeutic iron. The

crystalline form most commonly used for this purpose is the dihydrate, which contains 13.3% of crystallization water.

In this work, a NIRRS method based on the use of a fibre-optic probe for the determination of the water content in ferrous lactate dihydrate is proposed. Data are processed by using two different multivariate calibration techniques, viz. stepwise multiple linear regression (SMLR) and partial least-squares regression (PLSR). Because both are widely documented [9,10], only some general notions are given here.

Stepwise multiple linear regression performs a least-squares fitting of the reciprocal of the Lambert–Beer law. The concentration is modelled as a linear combination of the absorbance measured at different wavelengths in order to obtain the highest possible correlation for the following equation, with the smallest possible number of wavelengths:

$$C = k_0 + k_1A_1 + k_2A_2 + \dots + e \quad (3)$$

Partial least-squares regression is a variable-reduction method that compresses the information obtained throughout the spectrum into new variables called ‘PLS-components’ (PCs) that are subsequently used for calibration. The method simultaneously breaks down the absorbance and concentration matrices in such a way that the first PCs contain the greatest amount of information for predicting the sample values.

The value of each PC for each sample is called its ‘score’ and the regression coefficient for each variable at each factor its ‘loading’. These terms are related by the following equations:

$$\mathbf{X} = \mathbf{F}_X \mathbf{L}_X + \mathbf{E}_X \quad (4)$$

$$\mathbf{Y} = \mathbf{F}_Y \mathbf{L}_Y + \mathbf{E}_Y \quad (5)$$

where  $\mathbf{Y}$  denotes the concentration data matrix;  $\mathbf{X}$  the absorbance data matrix;  $\mathbf{F}_X$  and  $\mathbf{F}_Y$  the scores matrices;  $\mathbf{L}_X$  and  $\mathbf{L}_Y$  the loadings matrices; and  $\mathbf{E}_X$  and  $\mathbf{E}_Y$  the residuals matrices. The scores matrices are related by the following equation:

$$\mathbf{F}_Y = \mathbf{F}_X \mathbf{V} + \mathbf{E}_C \quad (6)$$

where  $\mathbf{V}$  is the internal relation and  $\mathbf{E}_C$  the residual matrix. The concentration of an unknown

sample,  $\mathbf{y}_0$ , is calculated from its spectrum,  $\mathbf{x}_0$ , by using the following expression:

$$\mathbf{y}_0 = \mathbf{x}_0(\mathbf{F}_Y^T \mathbf{X})^T \mathbf{V} \mathbf{L}_Y \quad (7)$$

After the quantitation conditions for the two methods are optimized, new samples can be determined directly in an almost instantaneous manner.

## 2. Experimental

### 2.1. Reagents

Karl Fischer methanol and formamide were obtained in RE and PRS grade, respectively, from Panreac; Karl Fischer reagent and Hidranal<sup>®</sup>-Composite 5 were purchased from Riedel de Haën; and ferrous lactate dihydrate ( $\text{C}_6\text{H}_{10}\text{O}_6\text{Fe} \cdot 2\text{H}_2\text{O}$ ) was supplied by Laboratorios Menarini.

### 2.2. Apparatus

The experimental set-up used consisted of an NIR-Systems 6500 near infrared spectrometer equipped with a reflectance detector and an AP6641 ANO4P fibre-optic module for qualitative and quantitative analyses. The instrument was operated via the bundled software (NSAS v. 3.20).

Karl Fischer titrations (KF) were carried out with a Metrohm 716 DMS Titrino autotitrator fitted to a 303 Ti Stand titration cell, also from Metrohm.

Spectral samples were processed by using the spectrophotometer's bundled software (NSAS), which includes the SMLR algorithm. The PLSR procedure was optimized with the aid of Unscrambler v. 503, from Camo (Trondheim, Norway).

### 2.3. Sample preparation

The stoichiometric content of water in ferrous lactate dihydrate is 13.3%. In order to ensure an adequate calibration range, samples with water contents over the range 11.4–14.6% were pre-

pared. Water contents exceeding that of the dihydrate were obtained by allowing ferrous lactate to stand in a moist environment for a few hours; on the other hand, water contents below that of the lactate were prepared by placing the product in a circular oven at 100°C under a nitrogen stream.

### 2.4. Determination of the water content by using the Karl Fischer method

The Karl Fischer method for the determination of the water content was implemented as follows: 40 ml of 75:25 (v/v) methanol/formamide mixture was placed in a cell and titrated with Karl Fischer reagent as far as the end-point indicated by polarized electrodes. An accurately weighed amount of about 0.1 g of ferrous lactate was then added to the cell and its contents titrated to a new end-point. Triplicate titrations for each sample were performed.

In order to check that iron did not interfere with the titration, one of the samples was analyzed thermogravimetrically in triplicate. The results obtained for that sample were  $13.61 \pm 0.02$  mg by the Karl Fischer method and  $13.69 \pm 0.17$  mg by the thermogravimetric method. The poorer precision of the latter can be ascribed to sample heterogeneity and the small amount used (about 20 mg).

### 2.5. NIRRS procedure

The spectrum of each sample was recorded in triplicate from 1100 to 2200 nm, the sample being turned over with a spatula between recordings. Spectra were averaged and subsequently used to obtain their first and second derivatives. Samples were measured in the vessel, simply by immersing the fibre-optic probe and squeezing it slightly against the sample. Fig. 1 shows the absorbance spectrum for one of the samples.

The results obtained in the quantitation of the samples were expressed as the relative standard error or prediction:

$$\text{RSEP}\% = \sqrt{\frac{\sum_{i=1}^n (C_{\text{KF}_i} - C_{\text{NIR}_i})^2}{\sum_{i=1}^n C_{\text{KF}_i}^2}} \times 100 \quad (8)$$

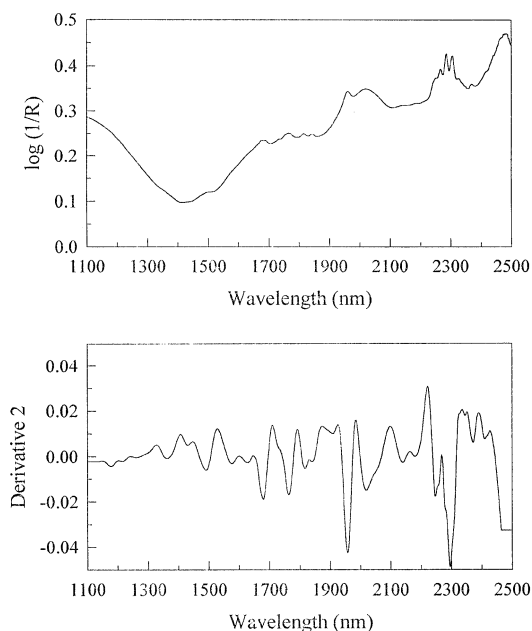


Fig. 1. NIRS spectrum of ferrous lactate dihydrate. Absorbance ( $\log 1/R$ ) and second derivative spectral modes.

where  $n$  is the number of samples, and  $C_{KF_i}$  and  $C_{NIR_i}$  are the water contents as determined by KF titration and the NIR method, respectively. RSEPC% and RSEPT% are used to refer to the relative standard error of prediction for samples

used in the calibration set and for those of the test set used to check the true predictive capacity of the model, respectively.

### 3. Results and discussion

The 22 samples available were split into two groups: 10 encompassing roughly uniformly the concentration range studied were used to construct the calibration set and the other 12 were employed as test set.

Table 1 gives the water contents of the samples as determined by the Karl Fischer method, as well as the standard deviations for the three determinations performed. The superscripts denote the samples that were used for calibration in the two methods (SMLR and PLSR).

The physical dependence of reflectance measurements results in spectral variability that should be minimized in order to ensure precise determinations of chemical components. Derivative spectra are known to decrease scattering effects and hence to increase the precision of quantitative analyses [11]. In this work, we used the absorbance and the first- and second-derivative spectral modes for comparison.

Table 1

Percentage of water in the samples assayed (determined by Karl Fischer titration)

Dried samples		Untreated samples		Moistened samples	
B1 <sup>a</sup>	11.14 ± 0.05	C1	13.22 ± 0.15	H1	14.11 ± 0.17
B2	11.24 ± 0.04	C2 <sup>a</sup>	13.36 ± 0.01	H2 <sup>a</sup>	14.15 ± 0.03
B3	11.27 ± 0.08	C3	13.46 ± 0.09	H3	14.28 ± 0.14
B4 <sup>a</sup>	11.37 ± 0.04	C4 <sup>a</sup>	13.86 ± 0.05	H4 <sup>a</sup>	14.58 ± 0.04
B5	11.98 ± 0.39				
B6	12.06 ± 0.06				
B7 <sup>a</sup>	12.06 ± 0.10				
B8 <sup>a</sup>	12.85 ± 0.27				
B9	12.97 ± 0.05				
B10	12.97 ± 0.06				
B11 <sup>a</sup>	13.05 ± 0.17				
B12	13.16 ± 0.02				
B13 <sup>a</sup>	13.19 ± 0.05				
B14	13.36 ± 0.05				

The confidence interval was calculated as the standard deviation for the three determinations performed.

<sup>a</sup> Indicates samples in the calibration set.

### 3.1. Determination by stepwise multiple linear regression

Appropriate selection of the measuring wavelengths is essential for calibration by SMLR. In this work, we used the ascending stepwise variant. This procedure involves using an initial wavelength resulting in the highest correlation between the absorbance and concentration to construct the calibration equation (Eq. (1) from a single term) and successively adding new wavelengths (terms of the calibration equation). The procedure can be adjusted by using wavelengths within preset ranges. Some authors [12] recommend using a wavelength in a region where the target analyte absorbs as the first measuring wavelength in order to ensure that the information obtained is due to the analyte rather than a product with which concentration it is correlated.

We compared the results obtained from wavelengths chosen according to three different criteria, namely:

- (A) Automatic selection of wavelengths throughout the spectrum (1100–2200 nm);
- (B) Selection of the first wavelength between 1350 and 1500 nm (around the water band at 1450 nm) and of the rest from any other region in the spectrum;
- (C) Selection of the first wavelength between 1900 and 2000 nm (around the water band at 1950 nm) and of the rest from any other spectral region.

The number of wavelengths to be used has to be optimized in order to avoid overfitting. This has been done in this work by applying an  $F$  statistical criterion to the summation of the squares of the residues ( $\text{SRC} = \sum e_i^2$ ) for the models to be compared. If a model involving  $n$  variables,

$$c_i = k_0 + k_1 A_1 + \dots + k_{g-1} A_{g-1} + k_g A_g + \dots + k_{n-1} A_{n-1} + e_i \quad (9)$$

is compared with that constructed from  $g$  variables,

$$c_i = k_0 + k_1 A_1 + \dots + k_{g-1} A_{g-1} + e_i \quad (10)$$

then the significance of the parameter set  $k_g, \dots, k_{n-1}$  must be proved.

The summation of the squares of the residuals for the reduced model ( $\text{SRC}_r$ ) and complete model ( $\text{SRC}_c$ ) can be calculated. The difference between the summation of the squares ( $\text{SRC}_{\text{exp}}$ ) arises from the presence of  $n - g$  variables in the complete model.

$F$  is calculated from

$$F = \frac{\frac{\text{SRC}_{\text{exp}}}{n - g}}{\frac{\text{SRC}_r}{m - n}} \quad (11)$$

where  $m$  is the number of samples used for calibration. If  $F_{\text{calc}} < F_{\text{tab}}$ , then some of the parameters in the series  $k_g, \dots, k_{n-1}$ , will be significantly non-zero.

Also, the goodness of fit can be estimated from the multiple correlation coefficient:

$$R = \sqrt{\frac{\sum_{i=1}^n (C_{\text{NIR}_i} - \bar{C}_{\text{KF}})^2}{\sum_{i=1}^n (C_{\text{KF}_i} - \bar{C}_{\text{KF}})^2}} \quad (12)$$

where  $\bar{C}_{\text{KF}}$  is the mean water content for the calibration sample as obtained with the Karl Fischer method. For this value to be comparable between the different models used, it must be corrected for the number of degrees of freedom,  $R_{aj}^2$ :

$$R_{aj}^2 = 1 - \left( \frac{m-1}{m-n} \right) (1 - R^2) \quad (13)$$

where  $n$  is the number of terms in the equation and  $m$  is as in Eq. (11).

While  $R^2$  increases with increasing number of terms and approaches unity,  $R_{aj}^2$  tends to level off once the optimum model is reached.

The figures of merit of the different calibration models ( $R_{aj}^2$ , RSEPC, and wavelengths used) are shown in Table 2. Results obtained from the first derivative were significantly worse, so they have been excluded from the table. As can be seen, criteria A and B produced very similar results, while C was slightly worse. The predictive capability of the best models (chosen in base of the greatest  $R_{aj}^2$  and the smallest RSEPC% values), both in absorbance and second derivative, was checked by predicting the test samples. The

Table 2

Relative standard error of prediction (%) found for samples of the calibration (RSEPC) and test set (RSEPT) by using SMLR

	Mode	$R_{aj}^2$	RSEPC (%)	RSEPT (%)	Wavelengths (nm)
Absorbance	A	0.995	0.7		2058, 1168
	B	0.995	0.5	1.4	1500, 1386, 1840
	C	0.988	0.8		2000, 1150
Derivative 2	A	0.997	0.3	1.3	2168, 1218, 1476
	B	0.995	0.5		1400, 1268, 2198
	C	0.991	0.6		1946, 1398, 1996

RSEPT (%) values found are also shown in Table 2 and were, as expected, very similar, the one obtained using second derivative spectra being slightly better.

### 3.2. Determination by partial least-squares regression

Partial least-squares regression enables the use of the whole spectral wavelength range while minimizing the problems arising from the high correlation between variables and measurement noise. However, with few samples, it may be advisable to reduce the number of variables by discarding those which contribute no information on the analyte. In this work, we assayed two different wavelength regions, namely the whole spectrum (1100–2200 nm) and the regions encompassing the two main absorption bands for water (1350–1550 and 1850–2000 nm). Both autoscaled data (variables centred and scaled to a unity variance) and unscaled data were used.

Models were constructed by cross-validation [13]. The number of significant PCs was taken to be the smallest for which the MSECv differed significantly from the minimum MSECv [14]:

$$\text{MSECv}_h = \frac{\sum_{i=1}^n (C_{\text{KF}_i} - C_{\text{NIR}_i})_h^2}{n} \quad (14)$$

where  $(C_{\text{KF}_i} - C_{\text{NIR}_i})_h$  is the difference between the concentrations obtained by the Karl Fischer method and by NIR for  $n$  samples, using  $h$  principal components to predict the PLS model.

Fig. 2 shows the variation of Mean Squared Error of prediction by Cross-Validation (MSECv) with the number of PCs for the models

tested. As can be seen, MSECv values were markedly smaller for absorbance data, which should therefore have a higher predictive power than first- and second-derivative data.

Table 3 gives the RSEPC (%) and RSEPT (%) values found for the different models assayed. The number of PCs used is given as a superscript in brackets in each case.

As expected, the RSEPC values obtained from absorbance data were the smallest. However the RSEPT (%) values were quite similar for the three spectral modes, ranging from 1.2 to 1.7%. When only the spectra corresponding to the water bands were used, the prediction errors were slightly smaller than those obtained from the whole spectrum; however, scaling the data made no difference.

Based on calibration errors, the model providing the best results was that constructed from the following parameters: spectral absorbance mode, wavelength range 1350–1550 and 1850–2000 nm, three PCs and unscaled spectra.

The SMLR and PLSR results were compared by obtaining the regression lines for the results provided by the best model in each case against the reference method, with both the calibration and the prediction samples. Table 4 shows the figures of merit for the lines. Note that the results obtained with both procedures were virtually identical and that the predictive capacity was quite good in both cases.

## 4. Conclusions

Use of the NIRRS technique in conjunction with a fibre-optic probe and a suitable calibration

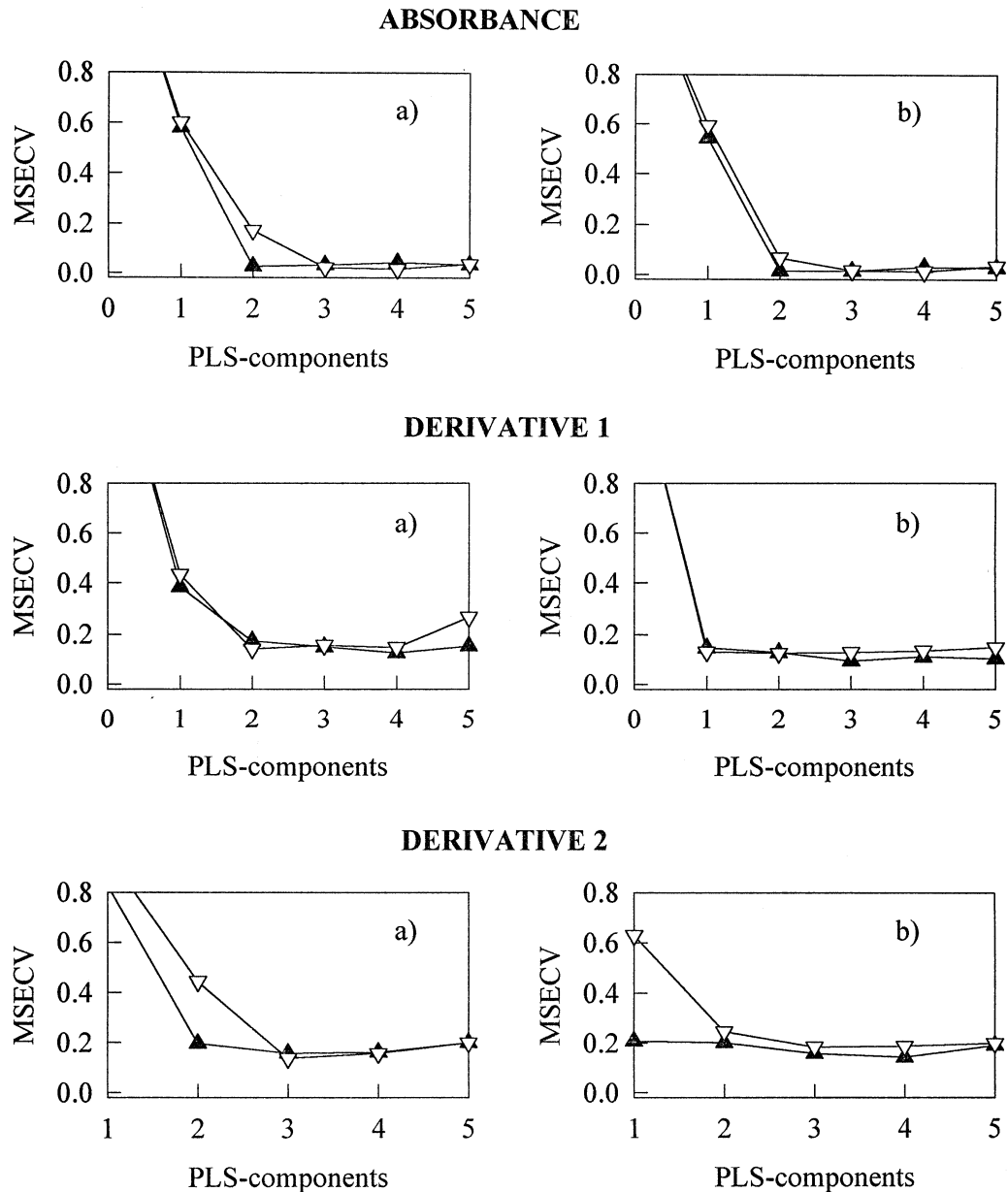


Fig. 2. Error of prediction in cross-validation for each PLS component in the different models and spectral modes tested. Wavelength ranges: (▲) 1100–2200 nm in absorbance and 1150–2200 nm in first and second derivative; (▽) 1350–1550 and 1850–2000 nm. (a) Unscaled data; (b) autoscaled data.

procedure enables non-destructive analysis of samples with good results as regards the determination of water in ferrous lactate. The method involves no sample pretreatment, so the analysis

time is substantially reduced, which is of practical significance for control analyses.

The calibration procedures tested, SMLR and PLSR, provide very similar results. Under the

Table 3

Relative standard error of prediction (%) for samples of the calibration (RSEPC) and test (RSEPT) obtained by the different PLSR models assayed

Range (nm)	Autoscaled		Not scaled	
	RSEPC (%)	RSEPT (%)	RSEPC (%)	RSEPT (%)
Absorbance				
1100–2200	0.7 <sup>(2)</sup>	1.5	0.9 <sup>(2)</sup>	1.6
1350–1550				
1850–2000	0.5 <sup>(3)</sup>	1.4	0.5 <sup>(3)</sup>	1.4
Derivative 1				
1150–2200	1.7 <sup>(2)</sup>	1.3	2.0 <sup>(2)</sup>	1.3
1350–1550				
1850–2000	0.8 <sup>(1)</sup>	1.3	1.9 <sup>(2)</sup>	1.3
Derivative 2				
1150–2200	2.8 <sup>(1)</sup>	1.7	1.6 <sup>(2)</sup>	1.0
1350–1550				
1850–2000	2.0 <sup>(3)</sup>	1.2	0.7 <sup>(2)</sup>	1.3

Superscripts in parentheses indicate the number of PLS components used in each model.

working conditions employed, which were aimed at developing a quality control procedure for raw materials, use of PLSR is recommended on account of its advantages for the detection of samples other than those in the calibration set. This increases the reliability of analyses [15,16].

### Acknowledgements

The authors are grateful to Spanish DGICyT for financial support granted for the realization of this work as part of Project SAF 93-0456.

Table 4

Figures of merit of the graphs provided by the best SMLR and PLSR models. Plot of concentration found by KF vs. NIR

	Calibration set	Test set
SMLR		
Intercept	0.0 ± 0.3	0.3 ± 1.6
Slope	1.00 ± 0.02	0.97 ± 0.12
<i>r</i>	0.999	0.983
PLSR		
Intercept	0.0 ± 0.4	0.4 ± 1.4
Slope	1.00 ± 0.03	0.97 ± 0.11
<i>r</i>	0.999	0.988

### References

- [1] The United States Pharmacopeia (USP 23), United States Pharmacopeial Convention, Rockville, MD, 1995.
- [2] HYDRANAL<sup>®</sup> Manual, Eugen Scholz Reagents, Riedel de Haën, Aktiengesellschaft, Seelze, Germany, 1988.
- [3] J.E. Sinsheimer, N.M. Poswalk, *Journal of Pharmaceutical Science* 57 (1968) 2007–2010.
- [4] C. Watson, *Analytical Chemistry* 49 (1977) 835A–840A.
- [5] J.K. Drennen, E.G. Kraemer, R.A. Lodder, *Critical Review in Analytical Chemistry* 22 (1991) 443–475.
- [6] K.M. Morisseau, C.T. Rhodes, *Drug Development and Industrial Pharmacy* 21 (1995) 1071–1090.
- [7] E.W. Ciurczak, J.K. Drennen, *Spectroscopy* 7 (1992) 12–14.
- [8] M. Blanco, J. Coello, H. Iturriaga, S. Maspocho, C. de la Pezuela, E. Russo, *Analytica Chimica Acta* 298 (1994) 183–191.
- [9] H. Martens, T. Naes, *Multivariate Calibration*, Wiley, Chichester, UK, 1989.
- [10] P. Geladi, B.R. Kowalski, *Analytica Chimica Acta* 185 (1986) 1–17.
- [11] L.S. Aucott, P.H. Gartjwaite, S.T. Buckland, *Analyst* 113 (1988) 1849–1854.
- [12] IQ2 Version 1.02 software manual, NIRSystems, Silver Spring, MD, 1990.
- [13] S. Wold, *Technometrics* 20 (1978) 397–406.
- [14] D.M. Haaland, E.V. Thomas, *Analytical Chemistry* 60 (1988) 1193–1202.
- [15] W. Linberg, J.Å. Persson, S. Wold, *Analytical Chemistry* 55 (1983) 643–648.
- [16] M. Blanco, J. Coello, H. Iturriaga, S. Maspocho, S. Alaoui-Ismaili, *Applied Spectroscopy* 50 (1996) 576–582.