

Influence of mathematical pretreatments on Iberian pig fat calibrations affected by uncontrolled error sources

V.M. Fernández-Cabanás,^a A. Garrido-Varo,^b J. García Olmo^b, E. De Pedro^b and P. Dardenne^c

^a*Escuela Universitaria de Ingeniería Técnica Agrícola. University of Seville. Ctra. Utrera Km. 1. 41013 Sevilla. Spain.*

^b*Escuela Técnica Superior de Ingenieros Agrónomos y de Montes. University of Córdoba. Apdo. 3048. 14080 Córdoba. Spain.*

^c*Centre de Recherches Agronomiques de Gembloux. Chaussée de Namur, 24. 5030 Gembloux. Belgium.*

Introduction

NIR analysis of fats and oils is strongly affected by uncontrolled variations related to the instrument, the environment, the sample preparation protocol, etc. It has been noted that the effect of these unexpected variations in the spectra are particularly important in the fat spectra versus those of other agro-food products, since while the latter generally present very broad absorption bands (more than 200 nm), fats and oils present very sharp and narrow peaks.¹

It has been demonstrated that the sensitivity of calibrations to the unexpected variation sources can be reduced by the use of repeatability files.² However, the implementation of repeatability files is not an easy task, as it requires a specific design for fats and oils.

The present paper tries to evaluate the effect of different mathematical pre-treatments on the sensitivity to unexpected sources of errors, of NIR equations developed for the prediction of fatty acids in pig fat.

Material and methods

Samples and reference data

341 samples of iberian pig fat with analytical data for the percentage of oleic, linoleic, palmitic and stearic fatty acids were used in this study. Fatty acid composition was obtained by gas chromatography. Global set was randomly divided into a calibration and a validation set. Chemical composition data for both sets is displayed in Table 1. Fat samples were analysed after melting in a microwave oven, and were maintained at 35 °C in an oven till they were scanned in a liquid state.³

NIRS hardware

Reflectance spectra were obtained on a Foss NIRSystems 6500 SY-I monochromator, from 400 to 2498 nm, every 2 nm. Analysis was performed using a spinning module. Samples were scanned on a transmittance cam-lock ring cell with 0.1 mm pathlength and provided with an aluminium reflectance surface (FOSS ref. IH-03459).

Table 1. Composition of calibration and validation sets

	Global Set				Calibration Set				Validation Set			
	C16:0	C18:0	C18:1	C18:2	C16:0	C18:0	C18:1	C18:2	C16:0	C18:0	C18:1	C18:2
N	341	341	341	341	311	311	311	311	30	30	30	30
Mean	21.08	10.67	52.32	9.39	21.07	10.64	52.36	9.38	21.20	10.96	51.90	9.45
Minim.	17.90	7.70	45.00	6.80	18.00	7.70	45.00	6.80	17.8	7.9	45.8	7.8
Maxim.	25.30	14.90	58.10	13.50	25.30	14.90	58.10	13.5	24.8	14.20	56.90	13.10
sd	1.46	1.32	2.44	1.32	1.45	1.31	2.43	1.33	1.64	1.46	2.48	1.27

Table 2. Combinations of pre-treatments used in calibration

Equation	Scatter correction	Derivative	Equation	Scatter correction	Derivative
1	None	None	29	SDT	1,10,10,1
2	SNV	None	30	SDT	2,5,5,1
3	DT	None	31	SDT	2,10,5,1
4	SDT	None	32	SDT	2,10,10,1
5	DTS	None	33	DTS	1,5,5,1
6	Standard MSC	None	34	DTS	1,10,5,1
7	Weighted MSC	None	35	DTS	1,10,10,1
8	Inverse MSC	None	36	DTS	2,5,5,1
9	None	1,5,5,1	37	DTS	2,10,5,1
10	None	1,10,5,1	38	DTS	2,10,10,1
11	None	1,10,10,1	39	Standard MSC	1,5,5,1
12	None	2,5,5,1	40	Standard MSC	1,10,5,1
13	None	2,10,5,1	41	Standard MSC	1,10,10,1
14	None	2,10,10,1	42	Standard MSC	2,5,5,1
15	SNV	1,5,5,1	43	Standard MSC	2,10,5,1
16	SNV	1,10,5,1	44	Standard MSC	2,10,10,1
17	SNV	1,10,10,1	45	Weighted MSC	1,5,5,1
18	SNV	2,5,5,1	46	Weighted MSC	1,10,5,1
19	SNV	2,10,5,1	47	Weighted MSC	1,10,10,1
20	SNV	2,10,10,1	48	Weighted MSC	2,5,5,1
21	DT	1,5,5,1	49	Weighted MSC	2,10,5,1
22	DT	1,10,5,1	50	Weighted MSC	2,10,10,1
23	DT	1,10,10,1	51	Inverse MSC	1,5,5,1
24	DT	2,5,5,1	52	Inverse MSC	1,10,5,1
25	DT	2,10,5,1	53	Inverse MSC	1,10,10,1
26	DT	2,10,10,1	54	Inverse MSC	2,5,5,1
27	SDT	1,5,5,1	55	Inverse MSC	2,10,5,1
28	SDT	1,10,5,1	56	Inverse MSC	2,10,10,1

NIRS software and chemometric treatments

All spectra were manipulated and processed, and all calibration equations were obtained, using ISI software NIRS3 ver. 4.0 and WINISI ver. 1.5 (Infrasoft International, Port Matilda, PA).

A total of 56 combination of pre-treatments (Table 2) were applied to obtain Modified Partial Least Squares (MPLS) calibration equations.^{4,5} The mathematical pre-treatments used were Standard Normal Variate (SNV), Detrending (DT),⁶ SNV and DT (SDT), DT and SNV (DTS), the three versions of Multiplicative Scatter Correction⁷ included in ISI software (Normal MSC, Weighted MSC and Inverse MSC) and four different derivative math treatments. The derivative math treatments are referred to by a four-digit notation (a,b,c,d).⁸ All calibrations were obtained for the spectral range 1100-2498 nm, for each chemical parameter, milling status, and a maximum number of outliers elimination passes of none, 2 or 9. Equations were obtained using an automatic routine (“*Teach automatic sequence*”) included in ISI software.

NIRS equations were evaluated, by examining the statistical values obtained for calibration and validation procedures. Main calibration statistics used were 1-VR (determination coefficient for cross validation) and SECV (standard error of cross validation), while validation was evaluated by SEP(C) (standard error of prediction corrected for bias) and R² (determination coefficient for validation).

Equation errors were compared using two test. First test sets, an arbitrary limit of 20% for significant differences with minimum error,⁹ and it is used for calibration and validation errors. Second test is based on a Fisher test,^{10,11} and defines a confidence interval (Equation 1) for errors with no significant differences with minimum error ($Error_{min}$)

$$(Error_{min}, Error_{min} \sqrt{F_{1-\alpha, n-1, n-1}}) \tag{1}$$

where α is the significance level (5% in this study) and (n-1) the degrees of freedom. This test was used to compare validation errors.

Results and discussion

Maximum and minimum cross validation errors (SECV) showed important variations for all parameters and number of outliers elimination passes (Table 3), as a result of the effect of the combination of the mathematical pre-treatments applied.

Table 3. Maximum, minimum and 20% limit for SECV values for calibrations with 0, 2 or 9 maximum outliers elimination passes. Grey background indicates significant differences found with minimum error.

	C16:0						C18:0					
Elim. passes	0		2		9		0		2		9	
	SECV	eqa	SECV	eqa	SECV	eqa	SECV	eqa	SECV	eqa	SECV	eqa
Minimum	0.38	5	0.27	15	0.18	39	0.32	5	0.25	16	0.22	7
Maximum	0.45	24	0.31	7	0.30	1	0.36	9	0.32	48	0.30	30
20% limit	0.46		0.32		0.22		0.38		0.30		0.26	
	C18:1						C18:2					
Elim. passes	0		2		9		0		2		9	
	SECV	eqa	SECV	eqa	SECV	eqa	SECV	eqa	SECV	eqa	SECV	eqa
Minimum	0.42	20	0.27	45	0.22	3	0.22	9	0.15	15	0.15	15
Maximum	0.45	1	0.32	1	0.28	1	0.28	7	0.20	1	0.20	7
20% limit	0.50		0.32		0.26		0.26		0.18		0.18	

There were not found significant differences for C16:0 and C18:1 using 0 and 2 passes of outliers elimination, while cross validation errors for C18:2 showed significant differences in all cases, indicating that the effect produced by the mathematical corrections depended to some measure on the fatty acid to be predicted.

Table 4. Maximum, minimum, 20% limit and F limit for SEP(C) values for calibrations with 0, 2 or 9 maximum outliers elimination passes. Grey background indicates significant differences found with minimum error.

Elim. passes	C16:0						C18:0					
	0		2		9		0		2		9	
	SEP(C)	eqa										
Minimum	0.35	19	0.39	14	0.37	20	0.27	37	0.26	5	0.26	1
Maximum	0.43	1	0.46	4	0.45	2	0.34	21	0.34	9	0.34	9
20% limit	0.42		0.47		0.44		0.32		0.31		0.31	
F limit	0.48		0.53		0.50		0.37		0.35		0.35	
Elim. passes	C18:1						C18:2					
	0		2		9		0		2		9	
	SEP(C)	eqa										
Minimum	0.57	5	0.58	21	0.57	28	0.27	48	0.23	36	0.27	48
Maximum	0.65	23	0.67	48	0.68	48	0.36	2	0.43	2	0.43	2
20% limit	0.68		0.70		0.68		0.32		0.28		0.32	
F limit	0.78		0.79		0.78		0.37		0.31		0.37	

Prediction errors SEP(C) for C18:1 didn't show significant differences, while calibrations for C16:0 and C18:0 only exhibited differences for 20% limit (Table 4). Differences for F test were only detected for C18:2 with outliers elimination passes. Table 4 also illustrates that range of variation among different passes of outliers elimination is reduced for SEP(C) values.

Conclusions

Despite that performance of prediction models varied with mathematical pre-treatment, general significant differences were not detected with the tests applied.

Optimisation of the derivatives used is has been found critical in calibrations for the prediction of fatty acids of pig fats.

External validation of calibrations developed using several combinations of pre-treatments is required for the optimisation of the spectral transformation (specially derivatives). The best pre-treatments in calibration and in prediction usually differ. So that, when making evaluation of pre-treatments it would be desirable to use some automatic software routines that permit an efficient evaluation of different validation sets and/or the execution of a large number of cross validation passes.

Unexpected spectral variations are not corrected with pre-treatments in calibration as well as using other strategies, like repeatability files.

Acknowledgements

This work was carried out using NIR hardware and software at the NIR/MIR Unit of the SCAI (University of Córdoba).

References

1. A. Garrido, J. García-Olmo, and M.D. Pérez-Marín, in *Near Infrared Spectroscopy in Agriculture. Monograph*, Ed by C. Robert, J. Workman and J. Reeves, ASA-CSSA-SSSA, Madison, WI, USA, (In press).
2. M.D. Pérez Marín, E. De Pedro, J. García-Olmo, and A. Garrido, in *Near Infrared Spectroscopy: The Future Waves*, Ed by A.M.C. Davies, and R.K. Cho, NIR Publications, Chichester, UK, p. 473, (2002).
3. J. García-Olmo, J., A. Garrido and E. De Pedro. 2001. *J. Near Infrared Spectrosc.* **9**, 49 (2001)
4. J.S. Shenk, and M.O. Westerhaus, *Crop. Sci.* **31**, 1694 (1991).
5. H. Martens and T. Naes, *Multivariate calibration*. John Wiley & Sons. Chichester, UK. (1989).
6. R.J. Barnes, M.S. Dhanoa and S.J. Lister, *Appl. Spectrosc.* **43**, 772 (1989).
7. H. Martens, S.A. Jensen and P. Geladí,. *Proceedings of the Nordic Symposium on Applied Statistics*, Stokkand Forlag Publishers, Stavanger, Norway. p 205 (1983).
8. J.S. Shenk M.O. and Westerhaus, *Routine operation, calibration, development and network system management manual*, NIRSystems, Inc., 12101 Tech Road, Silver Spring. MD 20904. USA. (1995).
9. W.R. Windham, D.R. Mertens and F.E. Barton II, in *Near Infrared Reflectance Spectroscopy (NIRS): Analysis of Forage Quality. Agriculture Handbook, N° 643*, Ed by G.C. Martens, J.S. Shenk and F.E. Barton II, USDA-ARS. US Government Printing Office, Washintong, D.C. p. 96 (1989).
10. T. Fearn, *NIR News.* **7(5)**, 5 (1996).
11. Y. Roggo, L. Duponchel, B. Noe and J.P. Huvenne, *J. Near Infrared Spectrosc.* **10**, 137 (2002).