Low-Field ^{1}H Nuclear Magnetic Resonance and Chemometrics Combined for Simultaneous Determination of Water, Oil, and Protein Contents in Oilseeds

Henrik Toft Pedersen, Lars Munck, and S¿ren Balling Engelsen*

Food Technology, The Royal Veterinary and Agricultural University, DK-1958 Frederiksberg C, Denmark

ABSTRACT: Prediction of the content of water, oil, and protein in rape and mustard seed was examined by a combination of low-field ^{1}H nuclear magnetic resonance (LF-NMR) and chemometrics, enabling utilization of the entire relaxation curves in the data evaluation. To increase the range of relative contents, the untreated seeds were wetted and dried; each treatment was followed by NMR analysis. The chemometric results are compared to traditional evaluation by multiexponential fitting of the relaxation curves. For this purpose, a new JackKnife validation procedure was developed to evaluate the number of exponential components objectively. Classification of the two kinds of seeds was easily performed by LF-NMR. Partial least squares regression to oil content in untreated rape and mustard seed yielded models with correlation coefficients of r = 0.88 and 0.89 with root mean square error of cross-validation (RMSECV) of 0.84 and 0.45, respectively. The rapeseed model was based on one component, whereas the mustard seed model was based on two components. If the seeds were dried, the predictive performance improved to r = 0.98 and RMSECV = 0.36 for rapeseed and to r = 0.95 and RMSECV = 0.38 for mustard seed. Upon drying, prediction of protein content in mustard seed improved, whereas the prediction of protein for rapeseed deteriorated. Global models, including the combination of untreated, wet, and dry seeds, all resulted in a robust and good predictive performance with RMSECV in the range 0.8–1.3% to water, oil, and protein content. It was demonstrated that drying the seeds to simultaneously determine water and oil content was not necessary when chemometrics was applied on the relaxation curves.

Paper no. J9500 in *JAOCS 77*, 1069–1076 (October 2000).

KEY WORDS: Chemometrics, 2D data, JackKnife, low-field NMR, mustard seed, oil, oilseeds, protein, rapeseed, water.

Rapid on-line measurements are of great interest in many industries, as they enable better control of production and quality. One method that shows promising results for fast on-line/at-line measurements, especially in the food industry, is low-field ^{1}H nuclear magnetic resonance spectroscopy (LF-NMR). The most common type of NMR is proton NMR, which is based on the fact that many foodstuffs are proton-rich, with protons originating, e.g., from water, fat, carbohydrates, and proteins. The advantages of LF-NMR, compared to other spectroscopic methods, are that it is noninvasive and nondestructive and that, perhaps most importantly, it is a bulk measure.

One of the early applications of LF-NMR was determination of fat content in different plant seeds by simply making a linear regression model that correlates the initial signal intensity or a ratio between two points (t_{11} and t_{70}) of the NMR free induction decay (FID) relaxation signal to the reference fat measurement (1). This procedure requires that the seed be dried before measurement, which is destructive as well as time-consuming. The next step in method development was to use a Hahn spin-echo (2) and perform a similar linear regression calibration based on a ratio between the initial amplitude and the spin-echo amplitude (3). The introduction of the spin-echo made it possible to determine the oil and water contents in oilseeds simultaneously; however, the samples still required a low moisture content (3–6). Gambhir (7) published a review of the application of LF-NMR to oilseeds, and a number of international standards have been published (8–10).

With new and fast bench-top NMR hardware it has become possible to acquire entire relaxation curves. A reasonable assumption would thus be that calibration based on the entire relaxation decay would reveal more information than calibrations based only on a single or a few points, particularly when attempting to estimate more components simultaneously, which is the aim of this work.

To handle the large amount of data points acquired, the traditional univariate linear or multiple linear regression methods no longer apply and, furthermore, LF-NMR relaxation data are highly co-linear (11), which is a problem for most analytical methods. Multivariate data analysis techniques, applied in chemistry since the 1970s under the name of chemometrics, include algorithms that easily handle both of the above-mentioned problems (12). A series of experiments were initiated to evaluate LF-NMR data with chemometrics for simultaneous deter-

*To whom correspondence should be addressed at Food Technology, The Royal Veterinary and Agricultural University, DK-1958 Frederiksberg C, Denmark. E-mail: se@kvl.dk

mination of multiple components in plant oilseeds. Tkachuk (13) used near-infrared reflectance (NIR) spectroscopy, including spectral information from several chemical bonds, to determine the amount of oil and protein in whole rapeseed. Velasco *et al.* (14) applied NIR to intact samples of Ethiopian mustard seed and were able to predict the fatty acid composition in the seed with good precision. However, resolution of molecular composition by NMR can only be achieved when using the more costly high-resolution NMR equipment. LF-NMR focuses on one major functional group, namely the protons, and their mobility. Now the question is whether LF-NMR can distinguish between different kinds of protons, e.g., those bound in free water or in more structured water, or even protons attached to oil, proteins, and carbohydrates for estimation of these fractions. Owing to its specificity in detecting protons, and its ability to measure bulk and not surface properties, LF-NMR could have an advantage compared to NIR.

MATERIALS AND METHODS

Samples and preparation. A sample set of 17 varieties of winter mustard and 20 varieties of spring rape was used for the experiment. Approximately 5 g of seed was weighed for each sample, and the weight was noted for later mass normalization of the acquired relaxation curves. After measurement of the untreated seeds, the seeds were stored in a desiccator over water for 10 d at 20°C for the seeds to absorb water. At this point, a new NMR measurement was performed, after which the seeds were dried overnight at 105°C, and a third NMR analysis was carried out. The contents of water, oil, and protein were calculated relative to sample mass for all three conditions to increase the relative range of contents. The NMR analysis was performed on the intact seeds at 25°C, and temperature equilibration before analysis was performed overnight in sealed NMR glass tubes.

NMR measurements. All NMR measurements were performed on a Maran bench-top pulsed NMR analyzer (Resonance Instruments, Witney, United Kingdom), equipped with a high-quality permanent magnet, operating at 23.2 MHz and with an 18-mm variable temperature probe head. The NMR equipment was controlled from a standard personal computer. An inversion recovery [INVREC (15)] experiment was performed where 22 "T_1-weights" (inversion delays: 0.1, 1, 10, 50, 60, 70, 80, 90, 100, 110, 120, 150, 200, 300, 400, 500, 600, 700, 850, 1000, 1200, and 1500 ms) were used. Furthermore, a FID-CPMG [Carr-Purcell-Meiboom-Gill (16,17)] pulse experiment was used, where the number of FID points is automatically calculated, depending on the selected tau and dwell time (distance between points). In this experiment, a tau value of 100 μs was used, allowing acquisition of 165 FID points with a dwell time of 0.5 μs acquired along with 4096 echoes in the CPMG part of the experiment. Only even-numbered echoes were used for subsequent data analysis. The sample relaxation delay was set to 2 s, and eight scans were accumulated for noise reduction.

Chemical reference data. The total water content of the seed was determined by drying the seeds overnight at 105°C. Replicates were made, and a replication error of 1 g/kg sample weight was accepted. The oil content was determined by the Soxhlet method (18,19). Samples of approximately 1 g were analyzed, and an error of replicates of 4 g/kg was accepted. Protein content was determined as nitrogen by Kjeldahl analysis (20) and converted to protein equivalents. Samples of approximately 1 g were analyzed, and an error of replicates of 2% was accepted.

Chemometrics. One of the main advantages of chemometric tools is that they are able to deal with spectral information that contains multivariate co-linear data, such as NMR relaxation data, by reducing data into a few functional or latent factors displayed in a graphical interface. In this work, principal component analysis (PCA) is applied to explore the origin of variation in the relaxation curves (11). Partial least squares regression (PLSR) is used to correlate the NMR relaxation data to chemical reference data, generating a multivariate linear regression model. For further in-depth descriptions of the use of PCA and PLSR we refer to other publications (21–23). All results reported are based on full cross (leave one out) validation (24). For reasons of comparison, global PLSR models that combined rape and mustard seeds were double-checked by performing cross-validation using two segments (rape and mustard), and global PLSR models that combined the three treatments were double-checked by performing cross-validation using three segments (untreated, wet, and dry). These results are generally not reported.

One of the few basic requirements for using PCA and PLS is that the data must be bi-linear. For NMR relaxation data this implies that, for a given time, the signal contributions from different protons must be additive, which in turn requires that instrument settings, such as receiver gain, must be kept fixed throughout the entire experiment and not optimally set for each individual sample.

Exponential fitting. To what extent exponential fitting can be used to describe compartmentalization in a sample, based on the characteristic relaxation times found in an exponential fit of the NMR relaxation data, has been treated elsewhere (25). Nevertheless, the amplitudes and relaxation time constants found in the exponential fit do describe some characteristics of the sample, and fitting is thus performed in this work to report changes in relaxation time constants in connection with the different treatments and to compare univariate data analysis with PLSR. For a further description of the fitting procedure, see Bechmann *et al.* (11).

Programs. All relaxation data were saved as binary quadrature data from the LF-NMR instrument. These files were imported to Matlab (The MathWorks Inc., Natick, MA) by an in-house-written routine, converted to magnitude format or phase-rotated as required, and then mass-normalized. The chemometric data analysis was performed with The Unscrambler (Camo, Trondheim, Norway). Exponential fitting was performed in Matlab with in-house software (11) available on www.models.kvl.dk.

RESULTS AND DISCUSSION

The range of contents of water, oil, and protein in intact samples

TABLE 1
Intercorrelations, *r*, and Contents of the Original Mustard Seed and Rapeseed Samples Used in the Experiment

	Untreated			Dry			Untreated + Wet + Dry		
Mustard seed	Water	Oil	Protein	Water	Oil	Protein	Water	Oil	Protein
Intercorrelations (*r*)									
Water	1	–0.64	0.47				1	–0.91	–0.95
Oil	–0.64	1	–0.63		1	–0.52	–0.91	1	0.80
Protein	0.47	–0.63	1		–0.52	1	–0.95	0.80	1
Content (%)	1								
Minimum	5.59	30.12	26.90		32	28.49	0.00	26.83	23.81
Maximum	6.12	33.70	28.92		35.7	30.7	14.60	32.00	28.49
Mean	5.83	31.63	28.02		33.57	29.77	7.00	31.22	27.68
Std. dev.	0.15	1.04	0.56		1.08	0.66	6.32	2.35	1.98
Rapeseed									
Intercorrelations (*r*)									
Water	1	–0.78	0.59				1	–0.84	–0.78
Oil	–0.78	1	–0.46		1	–0.70	–0.84	1.	0.45
Protein	0.59	–0.46	1		–0.70	1	–0.78	0.45	1
Content (%)									
Minimum	4.50	40.81	19.38		43.6	20.57	0.00	37.69	17.70
Maximum	5.65	48.17	23.05		50.7	24.22	13.93	50.70	24.22
Mean	5.03	44.76	20.80		47.16	21.92	5.96	44.33	20.61
Std. dev.	0.27	1.83	0.98		1.78	0.96	5.35	3.07	1.49

of the two kinds of seeds is not large, as shown in Table 1. This fact reduces the ability to model properly. Table 1 also lists intercorrelation among the three components as well as smallest and largest value, mean content, and standard deviation for reference data. The intercorrelation gives an indication of the extent to which a given model for, e.g., water is in fact based on the water-proton content or is confounded with some other component in the sample. A certain level of intercorrelation is expected (12) in biological samples, because water, oil, carbohydrates, and protein usually add up to approximately 100%.

Exponential fitting is performed on the acquired CPMG and INVREC data. Generally, determination of the number of relevant components poses a problem, as illustrated by Figure 1. The left side of the figure shows the raw CPMG data plus the under-

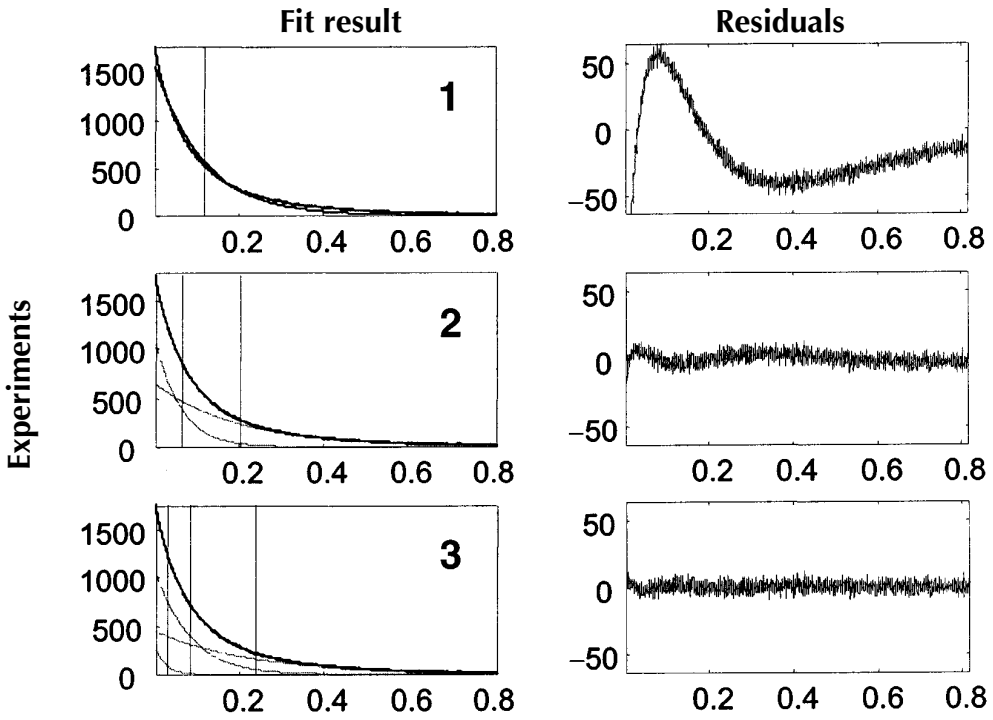

FIG. 1. Plot resulting from the exponential fitting routine. The left side shows the raw data plotted along with the underlying fitted exponential functions and indicators for the found time constants (vertical lines). The right side of the figure shows the residuals of the mono-, bi-, and tri-exponential fit.

lying fitted exponential functions extracted for one, two, and three exponentials. The right side of the figure shows the corresponding residuals, which usually form the basis for determining the appropriate number of components. Adding a new exponential component will always reduce the residual but with the risk that overfitting will occur. To avoid this potential pitfall, we have implemented a segmented cross-validation [or JackKnife resampling procedure (26)] routine for the exponential fitting procedure. The basic principle for the cross-validations is described in Figure 2, showing an example with only three segments. In this approach, a one-component curve fitting (calibration) is first performed on segments 2 and 3, and the root mean square (RMS) error E1 on segment 1 is recorded for validation. Then, the curve fitting is performed on segments 1 and 3, and the RMS error E2 on segment 2 is recorded for validation. Finally, the curve fitting is performed on segments 1 and 2, and the RMS error E3 on segment 3 is recorded for validation. By adding the RMS errors E1, E2, and E3, we now have a validated measure of the goodness of the fit to be compared with a corresponding two-component model. The downside of this approach is that it is quite time-consuming, as the number of fits needed to be calculated increases linearly with the number of segments used. However, it strongly improves the reliability of the selected number of components and makes the choice objective. Another way of determining the proper number of underlying latent factors is to perform PCA on the raw data. This gives a good indication of the dimension (rank) of the data set (matrix).

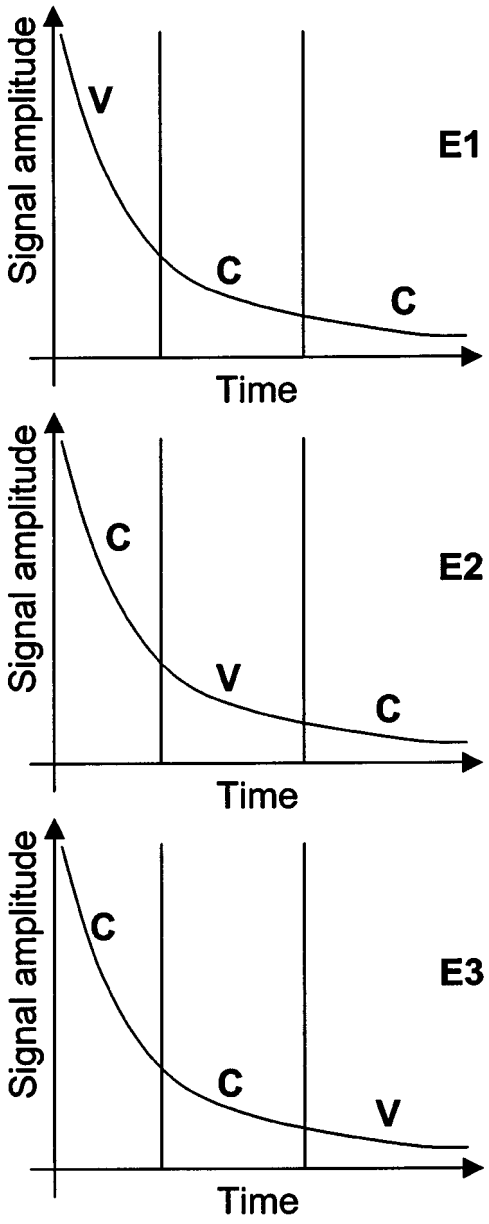

FIG. 2. Schematic description of segmented cross-validation routine used in the exponential fitting procedure. In this example three equal-sized segments are used. **E** is the summed error (**E** = **E1** + **E2** + **E3**) resulting from the individual steps; **C** = calibration; **V** = validation.

Figure 3 shows an example of the decrease in summed residual of fitted CPMG data for the rapeseed samples with the three different treatments. The figure shows clearly that the wet seeds require three components to properly describe the data when compared to both untreated and dry seeds, which require only two components. Three principal components were also necessary to describe the data variation in a PCA when the wet samples were included. This agreement between PCA and JackKnife validation of the exponential fitting procedure more than suggests that the water absorbed in the seeds does not entirely enter the normal compartmentalization of water in seeds, but introduces a new T_2 component. A similar JackKnife/PCA analysis was performed on the INVREC data and also consistently yielded two components for the wet samples. This result may indicate that we have fewer T_1 components than T_2 components or alternatively that we have more scarce information on the T_1 relaxation (4–6 parameters estimated from only 16 points). Table 2 reports an average of the obtained results. Both bi- and tri-exponential fitting was performed for all treatments, but only wet seeds required fitting of three underlying T_2 exponential curves. Cu-

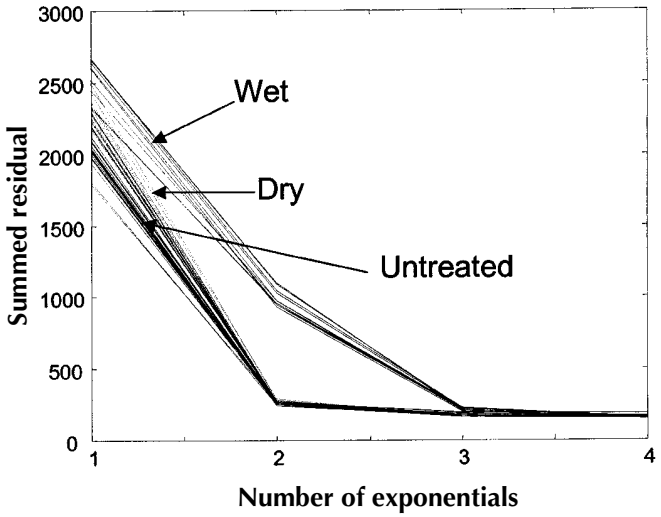

FIG. 3. Reduction in summed residual as a function of increasing number of exponentials fitted. Clearly the wet seeds require one more exponential in order to explain a sufficient level of the information.

TABLE 2
Mean Characteristic Time Constants for Bi- and Tri-exponential Fit of CPMG and INVREC Relaxation Data

	CPMG (ms)			INVREC (ms)	
	T_{20}	T_{21}	T_{22}	T_{11}	T_{12}
Mustard					
Dry	—	41	130	52	260
Untreated	—	40	130	39	180
Wet	2.3	44	140	35	140
Wet[a]	—	5.6	90		
Rape					
Dry	—	50	160	65	300
Untreated	—	50	160	45	210
Wet	3.0	55	170	39	170
Wet[a]	—	16	120		

[a]Two-component fits of wet seeds not justified by the JackKnife validation. CPMG, Carr-Purcell-Meiboom-Gill; INVREC, inversion recovery.

riously, the NMR relaxation of the wet seeds introduced a new short T_2 time constant denoted T_{20}, which is significantly faster than the time constants calculated in the two-component cases (dry and untreated). Emergence of the new short relaxation time constant upon adsorption of "very mobile" water/moisture indicates that part of the water does not penetrate into the inner structures of the seeds, but perhaps is in a state tightly bound to the surface structure of the seeds. A comparison of the two- versus three-component solutions for the wet seeds clearly shows that the additional short component is required to yield relaxation times, T_{21} and T_{22}, in the same range as for the dry and untreated seeds. This fact supports the validity of our JackKnife model, as the two original components persist; however, the physical origin of the short component remains to be definitively assigned.

Figure 4 shows what we call the exponential JackKnife plot [idea borrowed from Martens and Martens (27)], which results from a segmented cross-validation with eight segments. The plot shows the result of a bi-exponential fit of an untreated rapeseed sample where T_{21} (ms) is plotted against T_{21} (ms). The asterisks (covered by point number 5) from which all lines radiate show the relaxation times in the situation where the full relaxation profile is used and no validation is performed. Each of the other points (1 to 8) shows the situation where that segment has been left out and the fitting performed on the rest of the data. The plot shows that most information is stored in the initial part of the relaxation curve, because leaving out the first segment has the largest influence on the time constants found by the fit. The fact that most information is stored in the initial part of the relaxation curve can be explained by the nature of exponential functions, because for a given component they contribute the maximum to the initial part.

Figure 5A shows a global PCA based on INVREC, including both kinds of seeds and all three kinds of treatment. The first principal component (PC1) explains 90% of the variation

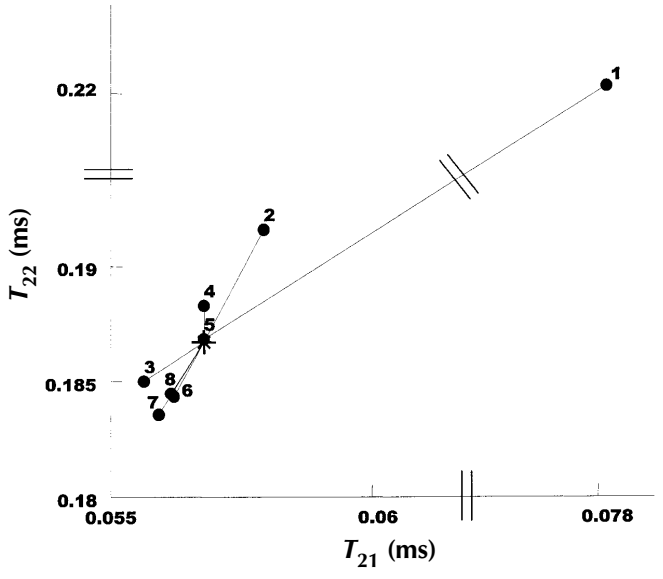

FIG. 4. JackKnife plot resulting from a segmented cross-validation routine when performing exponential fit. It can be seen that the majority of the information is stored in the first segment, resulting in the largest deviation from the full spectrum result if removed before fitting (thus used for validation). The asterisk under point #5 represents the position of the full spectrum fit.

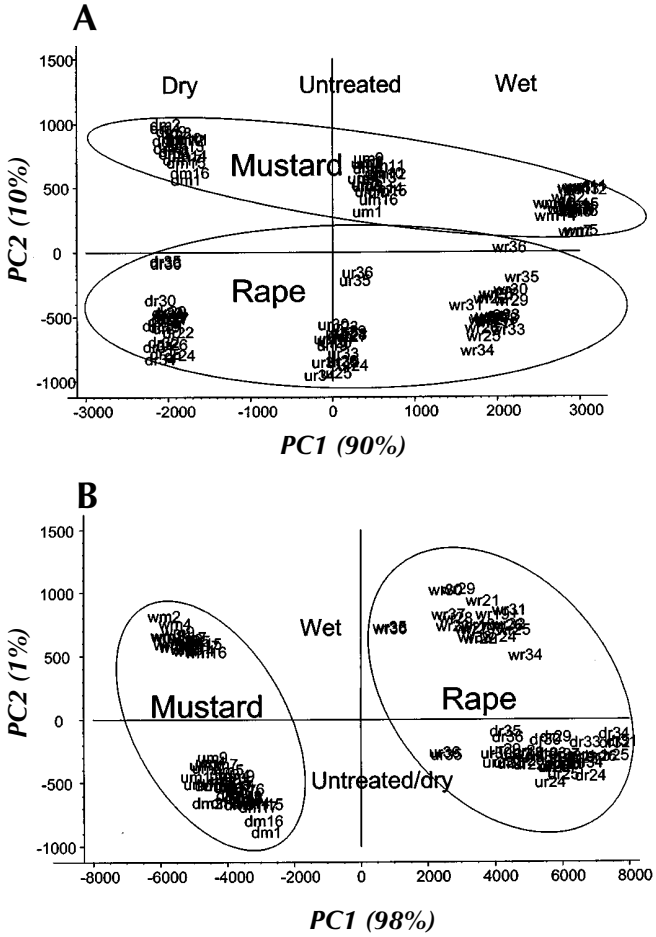

FIG. 5. Global principal component analysis based on (A) inversion recovery data and (B) Carr-Purcell-Meiboom-Gill data with both rape and mustard seeds and the three different kinds of treatments (untreated, wet, and dry) of the seeds. PC, principal component.

and describes the three different treatments, whereas the second principal component (PC2) explains the remaining 10% of the variation, including the information needed for seed classification. In Figure 5B, a global PCA for CPMG data is shown; however, in this analysis PC1 (98%) distinguishes the two kinds of seeds and PC2 (1%) mainly separates the wet seeds from the untreated and dry seeds, the latter being mixed up in one group. Including PC3 in the score plot does not enable further separation of the untreated and dry treatments. The difference in the information acquired by the different pulse experiments is quite interesting. It originates from the fact that INVREC probes differences in the matrix physics (spin-lattice relaxation), whereas CPMG probes differences in the protons' mobility (spin-spin relaxation). Thus, water and oil contents (not all oil is melted at the measurement temperature 25°C) are most likely well reflected in the CPMG data.

The easy separation of the two varieties of seeds and the different treatments implies that global models for distinguishing rape and mustard seed as a classification problem is quite simple. Looking at the relaxation times reported in Table 2, we can also clearly see the pattern of separation described above by the PCA analysis. The T_{21} relaxation times indicate especially that the water compartmentalization and mobility in mustard and rapeseed are quite different.

PLSR prediction models are shown in Table 3 for a combination of INVREC, FID, and CPMG data for untreated, dry and untreated, wet and dry together as well as global models (including both rape and mustard seeds). Almost all global models show good performance owing to the large range of contents in reference values. Prediction of oil content improves by drying the seed, whereas the protein prediction is ambiguous. Prediction models in which the three treatments are combined show good performance, probably partly owing to the enhanced range of contents. When using segmented cross-validation for the global models in the PLSR, as described in the Materials and Methods section, results comparable to the full cross-validation results were obtained. Generally, models for water and protein were slightly poorer, whereas models for oil generally showed improvements.

The prediction of protein in either untreated or dried seed is not satisfying. At present, it is not clear whether this fact is due only to the low variation or the relatively large standard deviation of the reference method. The fact that the models with the lowest correlation coefficients have the lowest root mean square error of cross-validation (RMSECV) indicates that variation in the reference material was not adequate. A new experiment with a larger protein variation should clarify this issue. We cannot, however, completely rule out the possibility that the good global protein calibration is due to intercorrelations with water and oil contents. However, in another application on water-holding capacity in frozen cod, we were able to predict water-holding capacity by using exactly the same approach with reasonable success ($r = 0.9$) in a sample material where there was no correlation between water-holding capacity and total water content. It is also indicated that the global model is not due to major intercorrelations (Table 1).

Figure 6A is a plot of measured vs. predicted oil content for a PLSR model of dried rapeseed using two components, and Figure 6B shows the segmented cross-validated results of the global model for all three treatments. In the latter, the validation is performed as segmented (the three treatments) cross-validation. By comparison with the full cross-validated results reported in Table 3, we observe an almost identical model with a slight improvement in the correlation coefficient, which indicates a sound model.

In conclusion, good prediction models were obtained for water and oil, and it also appears possible to make a prediction model for protein, if a sample material with more variation in protein content were collected. The results support the idea of using the complete spectral information acquired when analyzing a sample to perform the desired regression

TABLE 3
Full Cross-Validation of PLSR Models Using Combined INVREC, FID, and CPMG Data at 25¡C[a]

	Untreated			Dry			Untreated + wet + dry		
	r	RMSECV	#PC	*r*	RMSECV	#PC	*r*	RMSECV	#PC
Rape									
Water	0.90	0.12	1				0.98	1.00	2
Oil	0.88	0.84	1	0.98	0.36	2	0.97	0.77	2
Protein	0.81	0.57	3	0.73	0.61	1	0.86	0.72	2
Mustard									
Water	0.60	0.12	2				0.98	1.14	2
Oil	0.89	0.45	2	0.95	0.38	2	0.94	0.75	2
Protein	0.62	0.43	2	0.70	0.42	2	0.91	0.81	2
Global[b]									
Water	0.94	0.15	1				0.98	1.27	2
Oil	0.99	0.92	1	0.99	0.71	1	0.99	0.99	2
Protein	0.99	0.63	1	0.99	0.62	1	0.98	0.81	2

[a]#PC, number of partial least square regression (PLSR) components; *r*, correlation coefficient ; RMSECV, root mean square error of cross-validation; FID, free induction decay; for other abbreviations see Table 2.
[b]Describing separation of rape and mustard.

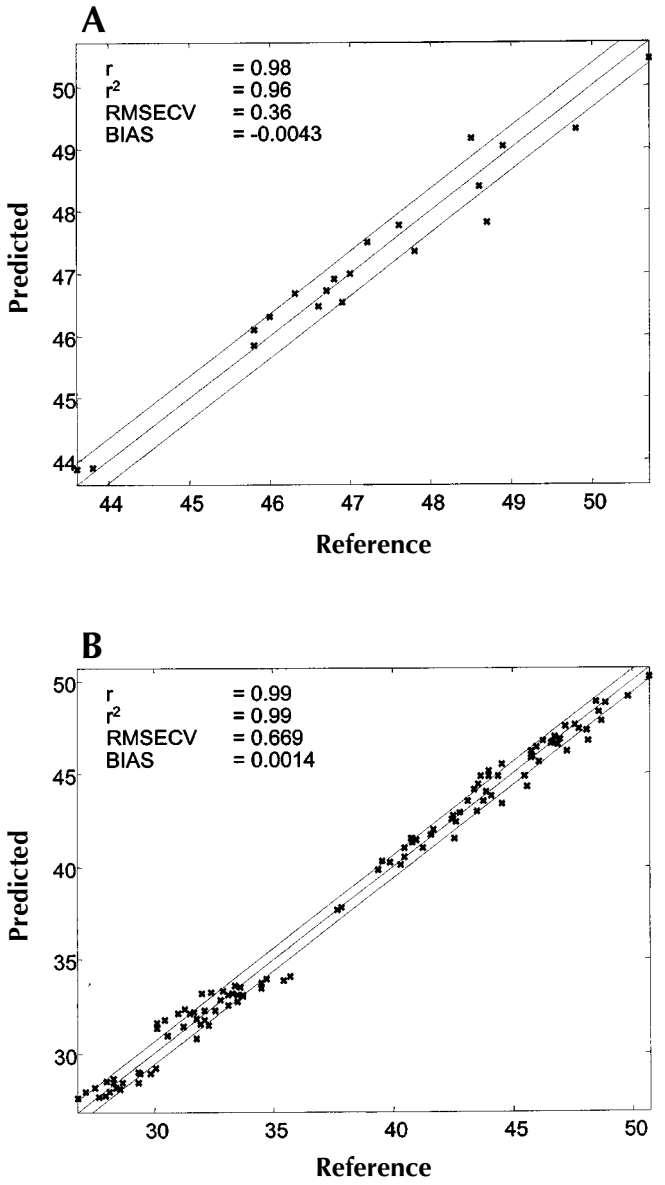

FIG. 6. Plot of predicted vs. measured oil content in dried rapeseeds for (A) full two components partial least squares regression (PLSR) and (B) two-component PLSR model based on both rape and mustard seed and all three treatments. This model is validated by segmented cross-validation.

instead of using just a single point or a few points, and they also demonstrate that simultaneous determination of several components is possible. The described multivariate methods used to analyze data from whole NMR relaxation curves should allow one to address the different classes of protons that represent water, oil, protein, and carbohydrates.

ACKNOWLEDGMENTS

We thank laboratory technician Kirsten Wilms for performing the chemical reference measurements and Gilda Kischinovsky for helpful editing of the manuscript. The samples and oil analyses were supplied by the cereal laboratory of Svalöf-Weibull AB (Svalöf, Sweden). The Danish Veterinary and Agricultural Research Council (SJVF) is acknowledged for salary to Henrik Toft Pedersen through the funding of The Centre for Critical Quality Attribute Determination in Muscle Foods.

REFERENCES

1. Tiwari, P.N., P.N. Gambhir, and T.S. Rajan, Rapid and Nondestructive Determination of Seed Oil by Pulsed Nuclear Magnetic Resonance Technique, *J. Am. Oil Chem. Soc. 51*:104–109 (1974).
2. Hahn, E.L., Spin Echoes, *Phys. Rev. 80*:580–594 (1950).
3. Rubel, G., Simultaneous Determination of Oil and Water Contents in Different Oilseeds by Pulsed Nuclear Magnetic Resonance, *J. Am. Oil Chem. Soc. 71*:1057–1062 (1994).
4. Tiwari, P.N., Pulsed NMR for Rapid and Nondestructive Determination of Oil in Oilseeds, *Bruker Minispec Application Notes 9* (1999) Bruker Analytische Messtechnik, Karlsruhe, Germany.
5. Srinivasan, V.T., Comparison of Different Pulse Sequences in the Nondestructive Estimation of Seed Oil by Pulsed Nuclear Magnetic Resonance Techniques, *J. Am. Oil Chem. Soc. 56*: 1000–1003 (1979).
6. Brosio, E., F. Conti, A. Nola, O. Scorano, and F. Balestrier, Simultaneous Determination of Oil and Water Content in Olive Husk by Pulsed Low Resolution Nuclear Magnetic Resonance, *J. Food Technol. 16*:629–636 (1981).
7. Gambhir, P.N., Application of Low-Resolution Pulsed NMR to the Determination of Oil and Moisture in Oilseeds, *Trends Food Sci. Technol. 3*:191–196 (1992).
8. AOCS, Simultaneous Determination of Oil and Moisture Contents of Oilseeds Using Pulsed Nuclear Magnetic Resonance Spectroscopy, *Official Methods and Recommended Practices of the American Oil Chemists' Society*, edited by D. Firestone, AOCS Press, Champaign, 1997, Recommended Practice Ak 4-95.
9. AOCS, Determination of Oil Content in Oilseeds, *Ibid.,* Official Method Am 2-93.
10. AOCS, Oil Content of Rapeseed by Nuclear Magnetic Resonance, *Ibid.*, Recommended Practice Ak 3-94.
11. Bechmann, I.E., H.T. Pedersen, L. Nørgaard, and S.B. Engelsen, Comparative Chemometric Analysis of Transverse Low-Field ^{1}H NMR Relaxation Data, in *Advances in Magnetic Resonance in Food Science,* edited by P.S. Belton, B.P. Hills, and G.A. Webb, Royal Society of Chemistry, London, 1999, pp. 217–225.
12. Jepsen, S.M., H.T. Pedersen, and S.B. Engelsen, Application of Chemometrics to Low-Field ^{1}H NMR Relaxation Data of Intact Fish Flesh, *J. Sci. Food Agric. 79*:1793–1802 (1999).
13. Tkachuk, R., Oil and Protein Analysis of Whole Rapeseed Kernels by Near Infrared Reflectance Spectroscopy, *J. Am. Oil Chem. Soc. 58*:819–822 (1981).
14. Velasco, L., J.M. Fernández-Martínez, and A. De Haro, Determination of the Fatty Acid Composition of the Oil in Intact-Seed Mustard by Near-Infrared Reflectance Spectroscopy, *Ibid. 74*: 1595–1602 (1997).
15. Vold, R.L., J.S. Waugh, M.P. Klein, and D.E. Phelps, Measurement of Spin Relaxation in Complex Systems, *J. Chem. Phys. 48*:3831–3832 (1968).
16. Carr, H.Y., and E.M. Purcell, Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments, *Phys. Rev. 94*:630–638 (1954).
17. Meiboom, S., and D. Gill, Modified Spin-Echo Method for Measuring Nuclear Relaxation Times, *Rev. Sci. Instrum. 29*:688–691 (1958).
18. Troëng, S., Oil Determination of Oilseed. Gravimetric Routine Method, *J. Am. Oil Chem. Soc. 32*:124–126 (1955).

19. Appelqvist, L.-Å., Further Studies on a Multisequential Method for Determination of Oil Content in Oilseeds, *Ibid. 44*:209–214 (1967).
20. AACC, Crude Protein—Kjeldahl Method, Boric Acid Modification, *Approved Methods of the American Association of Cereal Chemists*, American Association of Cereal Chemists, St. Paul, 1995, AACC Method 46-12.
21. Wold, S., K. Esbensen, and P. Geladi, Principal Components Analysis, *Chemom. Intell. Lab. Syst. 2*:37–52 (1987).
22. Geladi, P., and B.R. Kowalski, Partial Least-Squares Regression: A Tutorial, *Anal. Chim. Acta 185*:1–17 (1987).
23. Martens, H., The Philosophy of Partial Least Squares Regression, *Spectrosc. World 3*:26–27 (1991).
24. Martens, H., and P. Dardenne, Validation and Verification of Regression in Small Data Sets, *Chemom. Intell. Lab. Syst. 44*: 99–122 (1998).
25. Belton, P.S., Can Nuclear Magnetic Resonance Give Useful Information About the State of Water in Foodstuffs? *Comm. Agric. Food Chem. 2*:179–209 (1990).
26. Efron, B., *The Jackknife, the Bootstrap and Other Resampling*

Plans, Society for Industrial and Applied Mathematics, Philadelphia, PA (1982).
27. Martens, H., and M. Martens, Modified Jack-knife Estimation of Parameter Uncertainty in Bilinear Modelling (PLSR), *Food Qual. Preference 11*:5–16 (2000).

[Received January 3, 2000; accepted July 24, 2000]