Food Chemistry 121 (2010) 871–877

Contents lists available at ScienceDirect

Food Chemistry

journal homepage: www.elsevier.com/locate/foodchem

Analytical Methods

Development and validation of near infrared microscopy spectral libraries of ingredients in animal feed as a first step to adopting traceability and authenticity as guarantors of food safety

V. Fernández-Ibáñez [a], T. Fearn [b], A. Soldado [a], B. de la Roza-Delgado [a,*]

[a] Department of Animal Nutrition, Grasslands and Forages, Regional Institute for Research and Agro-Food Development, SERIDA, P.O. Box 13, 33300 Villaviciosa, Spain
[b] Department of Statistical Science, University College London, Gower Street, London WC1E 6BT, UK

ARTICLE INFO

Article history:
Received 19 February 2009
Received in revised form 9 September 2009
Accepted 13 October 2009

Keywords:
Near infrared (NIR) microscopy
Spectral libraries
Discriminant analysis
KNN
Safety
Ingredients
Animal feeds

ABSTRACT

Traceability of animal products has become a priority for governments of the developed countries as a guarantee of food safety. Near infrared microscopy (NIRM) has been proposed as an alternative technology to detect and quantify banned ingredients in feedstuffs. The great advantage of this technique is its objectivity, whilst retaining the sensitivity of classic microscopy. The aim of this work was to build an NIRM reference spectral library on animal feed, consisting of samples of animal feed ingredients and possible contaminants, and to assess its ability to discriminate between ingredients using an internal cross-validation. A total of 48,899 spectra were measured on 229 samples representing 30 different ingredients. The method chosen for classification was *K*-nearest-neighbours (KNN) using first derivative spectra. Although the results showed an overall classification error of 35.88%, there was good discrimination between ingredients of animal and vegetable origin. There was some confusion between similar vegetable ingredients but this is unimportant.

1. Introduction

Traceability of animals and animal products has become a priority for governments of the developed countries, due to consumer demand for comprehensive and integrated food safety policies. Therefore, the rendering industry and its customers have a great concern for the implementation of new analytical methods that ensure traceability and safety, allowing more informative labelling.

Directive 2002/2/EC says that "A detailed quantitative information may help to ensure that potentially contaminated feed materials can be traced to specific batches, which will be beneficial to public health and avoid the destruction of products which do not present a significant risk to public health". According to this directive the listing of feed materials for feeding stuffs shall be subject to the following rules for compound feeding stuffs intended for animals other than pets: listing of feed materials with an indication, in descending order, of the percentages by weight present in the compound feeding stuffs and as regards the above percentages, a tolerance of ±15% of the declared value shall be permitted.

In this context, the Official Journal of the European Union authorised method for the fight against fraud or accidental contaminations such as meat and bone meal is conventional optical microscopy. This technique enables identification of ingredients from vegetal (barley, straw, silage, wheat, etc.) or animal origin (muscle fibres, hair, feathers, bones, cartilage, etc.), which provides additional information on the type of meal involved (Gizzi & von Holst, 2002).

In recent years near infrared reflectance microscopy (NIRM) has been proposed as an alternative technology to detect and quantify banned ingredients in feedstuffs (de la Haba et al., 2007; de la Roza-Delgado et al., 2007a, 2007b; Piraux & Dardenne, 2000; von Holst et al., 2008). NIRM is an objective, sensitive and highly-selective technique; it combines the analytical advantages of microscopy and spectroscopy techniques (Baeten & Dardenne, 2001). The principle of NIRM is based on the collection of spectra from particles or small regions of a sample. These spectra can be collected from extremely small areas ($\leqslant$50 μm) using a Fourier transform near infrared reflectance (FT-NIR) instrument attached to a microscope with an optical system designed to focus on particles, increasing the efficiency of radiation transmission for microspectrum collection. The great advantages of this technique are that sample/particle recognition is not dependent on the expertise of the analyst and that it is possible to automate the procedure, increasing the number of samples analysed per unit of time compared with classical microscopy whilst retaining the sensitivity

* Corresponding author. Fax: +34 985890066.
E-mail address: broza@serida.org (B. de la Roza-Delgado).

0308-8146/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.foodchem.2009.10.072

advantage of microscopy. Summarising, the eyes and expertise of the microscopist are replaced by an infrared detector and a classification algorithm.

In order to develop a methodology for the complete screening of feedstuffs it is necessary to build NIRM reference spectral libraries with relevant sample information. It is well known that the development of a NIRM methodology to build spectral libraries is not a simple task, but tedious and time consuming. In previous works, Piraux and Dardenne (2000) and Baeten, Michotte Renier, Sinnaeve, and Dardenne (2001) have constructed spectral libraries including thousands of spectra of single particles from animal, vegetal and mineral feed ingredients and used these to identify the origin of unknown particles. The most important step of the process is to establish the experimental basis and parameters to be optimised: particle size, distribution, surface texture and number of spectra to average. These parameters were optimised in our previous research (Fernández-Ibáñez, Soldado, Vicente, Martínez-Fernández, & de la Roza-Delgado, 2008). All these parameters should be appropriately selected for the generation of large and robust spectral libraries including complete spectral information to characterise the ingredients included in animal feeds and capable of transfer between NIRM instruments. These experimental conditions represent the preliminary work for the development of identification methods with this methodology.

In animal feeds, due to real sample heterogeneity, several hundreds of particles must be analysed, the result of a sample analysis being a collection of hundreds of spectra. These spectra are the result of light absorption by organic molecules giving a "unique" fingerprint that can identify each compound (Burns, 1992; Clarke, Hammond, Jee, & Moffat, 2002). In this analysis by NIRM, the sample particles are spread on a sample holder and spectra taken from many small regions of the sample. Those ingredients included in the compound feed can then be identified by comparing these spectra with reference spectral libraries using an appropriate chemometric tool. This method has the advantage of relying only on the chemical properties of samples and being independent of human subjectivity.

Chemometricians have developed several useful methods for the analysis and interpretation of NIR spectra for the purpose of identifying and classifying the materials represented by the samples. NIRS users have begun to employ these techniques, variously called "qualitative analysis", "discriminant analysis", "product identification", "pattern recognition", or one of a number of lesser-used terms. All of these describe the same general concept, which is to determine from the NIR spectrum the nature of the sample (as opposed to determination of its composition). These advances are normally represented by computerised algorithms that allow the computer to be "trained" to "recognise" samples using a concept similar to that used in quantitative NIR analysis (Mark, 2001).

From the many possible approaches we have chosen to use the *K*-nearest-neighbours (KNN) algorithm (Næs, Isaksson, Fearn, & Davies, 2002). This classifies an unknown by finding its nearest (spectral) neighbours in the library. It is a standard nonparametric statistical technique with the advantages that it makes no distributional assumptions, is simple to apply, and would cope easily with any future expansion of the library.

The objective of this work was to build an NIRM reference spectral library consisting of samples of animal feed ingredients and possible contaminants, pretreated only by grinding, and to assess its ability to discriminate between ingredients using an internal cross-validation. Future work will study the external validation of the library, challenging it with mixed feed samples. These are the first steps towards providing a tool of practical use for replacing the current microscopy-based detection of ingredients and contaminants in animal feed.

2. Materials and methods

2.1. Samples

A spectral library was built using the most common ingredients included in feedstuffs together with banned ingredients such as processed animal proteins. The samples were provided by the largest feed industries and rendering plants in the north of Spain in the framework of project RTA2005-00212-C02-00 from the INIA (National Institute of Agro-food Research) from 2005 to 2008, and thus represent the variability encountered in the real production process. A total of 48,899 spectra were measured on 229 samples representing 30 different ingredients. All the ingredients included in this study and numbers of spectra collected are listed in Table 1.

The optimal methodology for sample pre-treatment and instrumental conditions to collect spectral data in this system has been investigated in previous work (Fernández-Ibáñez et al., 2008). In line with the findings there, the samples were ground to a particle size of 1 mm as the unique pre-treatment prior to NIRM analysis.

2.2. NIRM analysis

In this work an Auto Image Microscope connected to a Perkin–Elmer Spectrum One Fourier transform near infrared (FT-NIR) spectrometer in reflectance mode (1112–2500 nm) was used. This instrument enables the collection of spectra from small areas (50 μm × 50 μm) and it is equipped with an InGaAs detector, a camera and a viewing system to magnify the visible-light image. The latter may be used to localise and mark either individual particles or small areas to be analysed. In this work, the sample was spread on a sample holder as a continuous film, an area in the centre of the sample was selected and focussed on, and spectra were measured using fields of view of size 50 μm × 50 μm arranged in

Table 1
Ingredients, samples and total number of NIRM spectra included in the spectral library.

Ingredient	ns	nsp
Lucerne dehydrated	24	5109
Maize silage	23	4836
Fababean silage	2	450
Grass silage	13	2713
Grass hay	2	421
Cereal straw	27	5644
Beet pulp	14	3039
Cotton seed	7	1016
DDGS Barley	1	202
Soybean meal	15	3330
Sunflower seed	7	1578
Oats	4	847
Wheat	15	3210
Barley	21	4449
Rye	5	1146
Maize	21	4826
Bran	2	419
Corn flakes	1	210
Peas	2	427
Palm seed	1	218
Wheat middlings	1	210
Mineral correctors	6	1222
By-pass fat	3	630
Citrus pulp	3	663
Milk powder	1	234
Whey powder	1	210
Blood meal	1	235
Meat and bone meal	4	937
Haemoglobin	1	234
Animal plasma	1	234

ns: number of samples; nsp: number of spectra.

a 13 × 18 grid over this area, thus collecting approximately 200 spectra per sample. This method avoids any subjective selection of individual particles, while still representing the inherent variability in the sample. As in previous works in NIRM (Baeten et al., 2005; de la Roza-Delgado et al., 2007a) spectra were obtained from the ratio between raw spectra and the background (measured using the Spectralon plate), and the spectral information was stored as log (1/*R*), recorded at 4 nm intervals over the range 1112–2500 nm after conversion from cm^{-1} using the Perkin–Elmer software, Spectrum v. 5.01 (Perkin–Elmer Instruments, 2002). Each recorded spectrum was the average of 70 scans to reduce the noise in the spectral data (Fernández-Ibáñez et al., 2008).

2.3. Data treatment

In order to develop mathematical and statistical analysis, spectral data were exported from Spectrum v. 5.01 software in ASCII format into Matlab v7.4 (The Mathworks, Inc., Natick, MA, USA). Subsequent computations and chemometric analyses were carried out with programs written in Matlab by one of the authors (T.F.).

All models were developed in the spectral range of 1500–2448 nm since this is the range where noise is at a minimum. Moreover, it was decided to eliminate the main water peak, from 1824 to 1976 nm, to reduce the risk that differing moisture contents in the different ingredients should influence the classification rules.

Before trying to develop classification rules, the data were examined graphically, looking at spectra and at plots of principal component scores, to identify any obvious problems with particular spectra or samples.

2.4. Classification by KNN

The method chosen for classification was *K*-nearest-neighbours (KNN) (Næs et al., 2002). This is a nonparametric method that makes no assumptions about the shapes of the groups to be classified, making it particularly suited for use with data that are not multivariate normally distributed. It has the additional advantages of requiring very little tuning, and permitting subsequent expansion of the library without the need to refit models.

To classify an unknown sample we calculate its distance from each of the samples in the library, find the *K* nearest ones (*K* is a parameter to be tuned, in the range 1–25 in this work) and classify the unknown to the group that has the most members amongst these neighbours. Ties may be broken using some function of the distances themselves; here the average distance was used for this. This procedure is simple and intuitively appealing.

2.5. Validation

Since no model as such is built, it is easy to use cross-validation with KNN. The spectrum to be classified can simply be removed from the library used to classify it. To make the procedure more rigorous in this application, the entire sample to which a spectrum belonged was removed from the library used to classify it. This ensures independence between the unknown and the library. The exceptions to this were the ingredients represented by only one sample, for which only half the sample was removed.

2.6. Selecting an optimal pre-treatment

The KNN algorithm is based on spectral distances between samples, and these distances are affected considerably by the pre-treatments, if any, applied to the spectra. Thus an important part of tuning the algorithm is the selection of a pre-treatment.

The treatments compared here were no pre-treatment, detrends of orders, 0 (subtraction of a horizontal baseline) and 1 (subtraction of a sloping baseline) and derivatives of order 1 and 2, both computed using a Savitzky–Golay filter (The PLS Toolbox, version 3.5, Eigenvector Research, Manson, WA) with a window of 15 data points. In combination with some of these was applied the standard normal variate transformation (SNV) (Barnes, Dhanoa, & Lister, 1989, 1993). This method centres and scales individual spectra, correcting for variable scattering effects.

Another possibility that always improves speed and sometimes improves classification performance is to reduce the dimension of the spectra, compressing them to scores on a modest number of principal components. Here, compression to 10, 25 and 50 principal components was investigated.

3. Results and discussion

3.1. Removal of outliers

The preliminary inspection revealed two samples (one of barley, one of soya) with problems, both having spectra very different to other samples of the same ingredients. These were probably the result of either mislabeling or contamination and all the 443 spectra for these two samples were removed from the library, leaving 227 samples and 48,456 spectra.

Table 2
Discriminant performance for ingredients included in animal feeds. Mathematical treatments applied to spectral data.

Treatment	Minimum errors	Optimum *K*
No treatment	309	13
Detrend 0	239	9
Detrend 1	242	13
1 Derivative	248	18
2 Derivative	285	7

K: *K*-nearest-neighbours.

Table 3
Discriminant performance for ingredients included in animal feeds. PCA and SNV applied to spectral data.

Treatment	–			SNV		
	PC's	Minimum errors	Optimum *K*	PC's	Minimum errors	Optimum *K*
Detrend 0	–	239	9	–	247	22
	10	256	15	10	244	10
	25	273	10	25	245	6
	50	303	22	50	299	8
1 Derivative	–	248	18	–	270	21
	10	290	9	10	273	23
	25	280	10	25	267	12
	50	305	15	50	334	20

SNV: standard normal variate; PC's: principal components; *K*: *K*-nearest-neighbours.

Table 4
Results of leave-one-out cross-validation with 100% spectra of library with first derivative treatment and $K = 10$.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Lucerne dehydrated	1	3058	533	44	188	34	846	167	61	24	20	41	5	4	12	0	6	6	0	0	0	0	13	0	39	0	3	0	0	0	0
Maize silage	2	196	4132	5	73	8	229	13	9	2	4	3	15	10	79	13	35	1	1	1	1	0	1	0	5	0	0	0	0	0	0
Fababean silage	3	236	26	0	37	6	95	13	13	0	7	4	1	3	0	0	7	0	0	0	0	0	1	0	0	0	1	0	0	0	0
Grass silage	4	633	1084	24	469	35	360	23	46	2	9	6	3	2	1	0	2	0	0	0	0	0	7	0	6	0	1	0	0	0	0
Grass hay	5	193	40	5	35	2	113	2	1	0	6	9	3	0	3	0	0	0	0	0	0	0	5	0	0	0	4	0	0	0	0
Cereal straw	6	765	297	14	126	34	4099	163	23	1	10	5	25	4	22	0	0	3	0	1	0	1	40	0	6	0	0	0	0	0	5
Beet pulp	7	142	33	0	26	3	275	2444	4	3	1	4	19	13	38	2	8	6	0	0	0	2	1	0	9	0	1	0	0	0	0
Cotton seed	8	109	39	0	18	5	70	35	657	0	26	0	23	1	4	0	4	2	0	0	0	0	0	0	11	0	0	0	12	0	0
DDGS barley	9	22	9	0	0	0	2	1	1	144	6	16	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Soybean meal	10	43	3	1	0	4	4	5	85	1	2690	69	4	4	7	1	13	3	0	8	0	1	3	0	0	0	16	0	155	0	0
Sunflower seed	11	106	15	5	3	7	21	10	3	6	168	1228	0	0	1	0	1	0	0	0	0	0	0	0	0	0	4	0	0	0	0
Oats	12	5	170	0	2	1	32	50	11	1	0	1	150	7	242	19	129	8	0	0	0	9	5	0	2	0	3	0	0	0	0
Wheat	13	2	8	0	2	0	2	3	2	0	2	0	11	1783	1027	108	225	7	3	20	0	0	5	0	0	0	0	0	0	0	0
Barley	14	3	98	0	3	2	69	18	1	1	8	1	111	251	2698	404	463	19	0	34	0	2	27	0	0	0	3	0	0	0	0
Rye	15	1	7	0	0	0	3	2	0	0	1	0	13	58	775	169	111	0	0	3	0	0	2	0	0	0	1	0	0	0	0
Maize	16	2	21	0	0	0	5	7	0	0	5	0	26	477	865	222	3165	7	1	14	0	0	8	1	0	0	0	0	0	0	0
Bran	17	20	21	0	2	0	5	16	7	4	8	0	65	9	100	2	69	53	2	0	5	29	1	0	1	0	0	0	0	0	0
Corn flakes	18	0	0	0	0	0	0	0	0	0	0	0	1	25	61	7	116	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Peas	19	0	0	0	0	1	0	0	0	0	1	0	0	23	125	2	15	0	0	229	0	0	24	0	0	0	7	0	0	0	0
Palm seed	20	6	0	0	0	0	2	9	0	1	1	1	1	0	1	0	2	0	0	0	194	0	0	0	0	0	0	0	0	0	0
Wheat middlings	21	4	9	0	0	0	14	3	0	0	0	0	23	2	32	3	16	18	1	0	0	77	6	0	0	0	2	0	0	0	0
Mineral correctors	22	33	3	0	1	0	15	2	0	1	6	10	6	9	122	4	9	4	0	31	0	4	942	16	0	0	4	0	0	0	0
By-pass fat	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	630	0	0	0	0	0	0	0
Citrus pulp	24	27	25	1	10	0	4	101	1	0	0	0	5	1	8	6	1	0	0	1	0	0	2	0	239	231	0	0	0	0	0
Milk powder	25	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	231	0	0	0	0	0	0
Whey powder	26	1	1	0	0	0	0	0	0	0	8	3	0	0	0	0	0	0	0	1	0	0	0	0	1	0	195	0	0	0	0
Blood meal	27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	235	0	0	0
Meat and bone meal	28	0	0	0	0	0	0	0	2	0	10	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	924	0	0
Haemoglobin	29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	234	0
Animal plasma	30	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	232

3.2. Tuning the classification algorithm

Because the number of combinations studied was large and a full cross-validation takes several hours to compute, the initial comparison of pre-treatments was carried out using a restricted cross-validation. For each pre-treatment option a 1% stratified sample, comprising 675 spectra balanced across ingredients, was classified by comparison with the library. As described in Section 2, the sample to which the spectrum belonged was withheld from the library used to classify it. A second short cut, again with the aim of saving computing time, was that the principal component analysis was carried out once only, on the whole library. With so many spectra this is unlikely to have had any serious effect on the cross-validation. All the treatments were run with a range of K from 1 to 25. The results in Tables 2 and 3 are the best ones obtained for each treatment.

The first step was to compare the basic pre-treatments of detrend and derivative. The results are given in Table 2. With no mathematical pre-treatment the number of incorrectly assigned spectra was 309, with a K of 13. This is poorer performance than that obtained with the use of any of the detrend or derivative treatments. Comparing the treatments, there seems little to choose between the two detrend options and first derivative, all three of these being preferable to second derivative. The two options of a detrend of order 0 and a first derivative were carried forward to the next stage, and the others discarded.

In this second stage, the use of dimension reduction by principal component analysis and the application of SNV was combined with the two retained options. Table 3 shows the results. It appears that neither the dimension reduction nor the SNV treatment improves the error rate of the classification. In applying the principal component analysis the scores were scaled to have equal variances, rather than variances proportional to their eigenvalues, which is why 50 scores (which capture nearly all of the variability in the original spectra) can give a result very different to the original one. KNN is sensitive to this rescaling of dimensions.

Since the extra sophistication of the methods studied in Table 3 brings no benefit, we are left to choose between the three equally good options in Table 2, the two detrends and the first derivative. All of these give quite a high error rate, in the region of 35%, though it will be seen later that this is not as bad as it seems. From the three options, first derivative was chosen. One motivation for this was that the derivative tends to emphasise some of the spectral features, making interpretation a little easier than with the raw spectra.

3.3. Selection of number of neighbours, K

Having fixed the pre-treatment a full cross-validation was run for each value of K in the range 1–25 and a value of $K = 10$ chosen as that which gave the best results. In fact the number of errors was not strongly dependent on the value of K, and any choice in the range 7–15 would give similar results. Table 4 shows the results of this full cross-validation for the chosen value $K = 10$. Table 4 has the form of a confusion matrix, showing the number of correct classifications for each ingredient, as well as a cross-tabulation of all the misclassifications. For example 538 spectra from lucerne were misclassified as maize silage. The average classification error was 35.88%, very similar to that obtained with the 1% sample.

4. Discussion of results

Although the total number of misclassifications is large, the confusions are concentrated in a limited number of areas, and many of them are unimportant.

It is important to remark that there is not much confusion between ingredients of animal and vegetable origin, although there is a lot of confusion between similar vegetable ingredients. A basic similarity between the ingredients and a wide variability between individual spectra of the same ingredient has induced many classification errors between spectra within the family of forages. This could be due to the fact that they have similar cell wall structures. Their nutritive value is usually assigned to the aerial part of the plant, which is related to the distribution of resources into cell contents. All the most relevant spectral similarities between plant tissues are related to cellulose, especially around 1700, 1900, 2100,

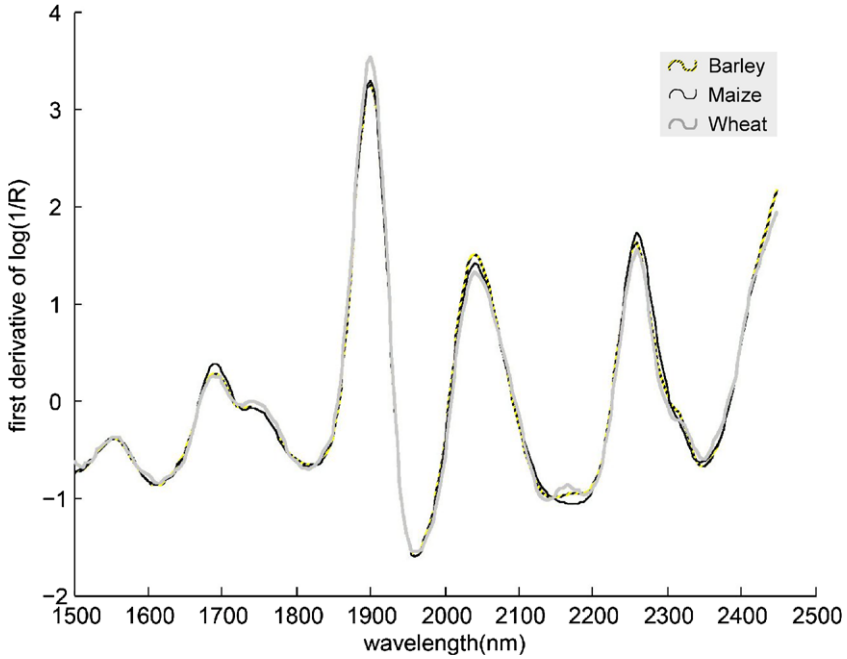

Fig. 1. First derivative spectral characteristics of barley, maize and wheat average spectra.

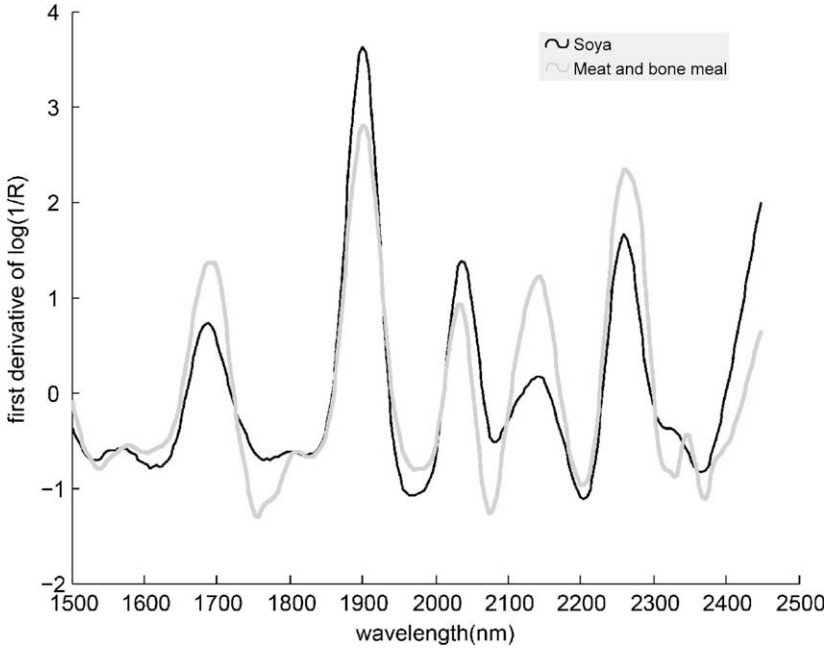

Fig. 2. First derivative spectral characteristics of soybean and meat and bone meal average spectra.

2250 and 2350 nm; lignified structures are related with 1685 nm, among others (Shenk,Workman, & Westerhaus, 2008).

Looking at the cereals group, we find that the majority are correctly classified, although there are many internal confusions, the most important ones being between barley, maize and wheat. These mistakes are reasonable, both spectral shape and nutritive values are quite similar, especially as regards starch content. Fig. 1 shows the first derivative mean NIRM spectra of barley, maize and wheat. The three spectra present similar absorption bands at 1510–1540 nm related to N–H and O–H stretch first overtone, bands associated with protein (1510 nm) and starch (1540 nm); at 1780–1820 nm related to C–H stretch first overtone, a band associated with cellulose; at 2055–2060 nm, a band associated with protein; and at 2270–2300 nm, related to second overtone protein, C–O and O–H stretch combination cellulose (2270 nm) and C–H stretch/CH_2 deformation starch (2280 nm) (Shenk et al., 2008). Discrimination between spectra is possible at around 2140–2190 nm, corresponding to a maximum for wheat, perhaps due to the higher content of bran compared with maize and barley (de Blas, Mateos, & Rebollar, 1999). These clearly distinct spectra are of course averages of thousands of individual spectra. It is the variability within samples that causes the misclassifications.

A more important confusion is that between soybean meal, of vegetable origin, and meat and bone meal (MBM), of animal origin. From Table 4 we see that 155 soybean spectra are classified as MBM, corresponding to 5% of the total soybean spectra included in the library. However, only 1% (10 spectra) of MBM are classified as soybean. Fig. 2 shows the mean spectra obtained from the soybean and MBM samples. The most striking differences are observed in bands in the region between 2300 and 2400 nm, mostly attributed to the fat content of the MBM. Again it needs to be remembered that although the mean spectra are clearly different, this does not mean that all the individual spectra will be. The confusions may arise from the high protein levels in both ingredients, corresponding to the following characteristic wavelengths: 1510, 2055–2060, 2180 and 2300 nm. von Holst et al. (2008) identify MBM traces and achieve the differentiation of animal from vegetable feed ingredients based on the evaluation of near-infrared spectra obtained from individual particles in sedimented samples. They develop decision rules using the absorbances at a small number of specific wavelengths. It is possible that the confusions seen here between soybean meal and MBM spectra collected directly on ground samples could be avoided by selecting specific NIR regions.

5. Conclusions and future work

Near infrared microscopy spectral libraries using appropriate chemometric tools can provide an objective technique to identify the ingredients included in compound feeds by comparing spectral information from the feed with the reference spectral library. Previous studies, cited in the introduction, have focussed on the use of sedimented samples. In contrast, the samples studied here have been pretreated only by grinding.

The results using full cross-validation and a value of $K = 10$ on first derivative spectra have shown that there is not much confusion between ingredients of animal and vegetable origin. Although there is some confusion between similar vegetable ingredients this is unimportant in the content of detecting contamination. It is the similarity of the ingredients and the variability within samples that causes these misclassifications.

Although for complete identification the method showed some limitations, this study opens up an alternative way to use NIRM spectral libraries of ingredients of animal feed as a screening tool for checking feed composition, without the need for sedimentation of the samples. Much more work is needed before NIRM can replace the current microscopy for an infrared detector and a classification algorithm. In particular, future work will deal with the external validation of the library, challenging it with mixed feed samples. However, the findings reported here represent an encouraging first step towards this goal.

Acknowledgements

This research work was developed during a stay of the first author in the Department of Statistical Science of University College London with Prof. Tom Fearn, with a grant from INIA (National Institute of Agro-food Research). The work was supported by the

Spanish project RTA2005-00212-C02-00 from the INIA. The authors are grateful to INIA for a pre-doctoral fellowship of Mª del Valle Fernández-Ibáñez and post-doctoral financial support of Dr. Ana Soldado. Thanks are also given to Alfonso Carballal and Nutrition Laboratory staff for their technical assistance.

References

Baeten, V., & Dardenne, P. (2001). The contribution of near infrared spectroscopy to the fight against the mad cow epidemic. *NIR News, 12*(6), 12–13.

Baeten, V., Michotte Renier, A., Sinnaeve, G., & Dardenne, P. (2001). Analyses of feedingstuffs by near-infrared microscopy (NIRM): Detection and quantification of meat and bone meal (MBM). In *Proceedings of the sixth international symposium on food authenticity and safety, Nantes, 28–30 November* (pp. 1–11).

Baeten, V., Von Holst, C., Garrido, A., Vancutsem, J., Michotte Renier, A., & Dardenne, P. (2005). Detection of banned meat and bone meal in feedstuffs by near-infrared microscopic analysis of dense sediment fraction. *Analytical and Bioanalytical Chemistry, 382*(1), 149–157.

Barnes, R. J., Dhanoa, M. S., & Lister, S. J. (1989). Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy, 43*, 772–777.

Barnes, R. J., Dhanoa, M. S., & Lister, S. J. (1993). Correction to the description of standard normal variate (SNV) and de-trend (DT) transformations in practical spectroscopy with applications in food and average analysis – 2nd edition. *Journal of Near Infrared Spectroscopy, 1*(3), 185–186.

Burns, D. A. (1992). In D. A. Burns & E. W. Ciurczak (Eds.), *Handbook of near-infrared analysis* (pp. 1–5). New York: Marcel Dekker.

Clarke, F. C., Hammond, S. V., Jee, R. D., & Moffat, A. C. (2002). Determination of the information depth and sample size for the analysis of pharmaceutical materials using reflectance near-infrared microscopy. *Applied Spectroscopy, 56*(11), 1475–1483.

de Blas, C., Mateos, G. C., & Rebollar, P. G. (1999). FEDNA para la formulación de piensos compuestos. In *Fundación Española para el Desarrollo de la Nutrición Animal*. Madrid: Ediciones Peninsular.

de la Haba, M. J., Fernández-Pierna, J. A., Fumière, O., Garrido-Varo, A., Guerrero, J. E., Pérez-Marín, D. C., et al. (2007). Discrimination of fish bones from other animal bones in the sedimented fraction of compound feeds by near infrared microscopy. *Journal of Near Infrared Spectroscopy, 15*, 81–88.

de la Roza-Delgado, B., Soldado, A., Martinez-Fernández, A., Vicente, F., Garrido-Varo, A., Pérez -Marín, D., et al. (2007a). Application of near-infrared microscopy (NIRM) for the detection of meat and bone meals in animal feeds: A tool for food and feed safety. *Food Chemistry, 105*, 1164–1170.

de la Roza-Delgado, B., Soldado, A., Modroño, S., Martinez-Fernández, A., Vicente, F., Garrido-Varo, A., et al. (2007b). Near infrared spectroscopy and quantify animal meals in feedstuffs: Are real certified samples necessary to help validate the results? In G. R. Burling-Claridge, S. E. Holroyd, & R. M. W. Sumner (Eds.), *Near infrared spectroscopy: Proceedings of the 12th international conference* (pp. 140–145). Chichester, West Sussex, UK: NIR Publications.

Fernández-Ibáñez, M. V., Soldado, A., Vicente, F., Martínez-Fernández, A., & de la Roza-Delgado, B. (2008). Particle size optimisation in development of near infrared microscopy methodology to build spectral libraries of animal feeds. *Journal of Near Infrared Spectroscopy, 16*, 43–48.

Gizzi, G., & von Holst, C. (2002). *Rendering processes and the detection of meat and bone meal in animal feed*. Italy: Joint Research Centre of the European Commission, Institute for Health and Consumer Protection, Food Products Unit Ispra.

Mark, H. (2001). Qualitative near-infrared analysis. In P. Williams & K. Norris (Eds.), *Near infrared technology in the agricultural and food industries* (pp. 233–238). St. Paul, MN, USA: American Association of Cereal Chemists.

Næs, T., Isaksson, T., Fearn, T., & Davies, T. (2002). *Multivariate calibration and classification*. Chichester, West Sussex, UK: NIR Publications. pp. 221–259.

Perkin–Elmer Instruments LLC (2002). IR spectroscopy software: User's guide. UK.

Piraux, F., & Dardenne, P. (2000). Feed authentication by near-infrared microscopy. In A. M. C. Davies & R. Giangiacomo (Eds.), *Near infrared spectroscopy: Proceedings of the ninth international conference* (pp. 535–541). Chichester, West Sussex, UK: NIR Publications.

Shenk, J., Workman, J. J., Jr., & Westerhaus, M. O. (2008). Application of NIR spectroscopy to agricultural products. In A. Burns & E. W. Cuirczak (Eds.), *Handbook of near-infrared analysis* (3rd ed.. *Practical spectroscopy series*, pp. 347–386). New York: Marcel Dekker.

von Holst, C., Baeten, V., Boix, A., Slowikowski, B., Fernámdez-Pierna, J. A., Tirendi, S., et al. (2008). Transferability study of a near-infrared microscopic method for the detection of banned meat and bone meal in feedingstuffs. *Analytical and Bioanalytical Chemistry, 392*, 313–317.