MACROSCOPIC NEAR-INFRARED REFLECTANCE SPECTROSCOPY (WP5)

I. Murray[1], A. Garrido-Varo[2], D. Perez-Marin[2], J.E. Guerrero[2], V. Baeten[3], P. Dardenne[3], S. Termes[4], J. Zegers[5] and R. Frankhuizen[6]

[1] SAC, Scottish Agricultural College, Aberdeen, UK.

[2] UCO, Department of Animal Production, Faculty of Agriculture and Forestry Engineering. University of Cordoba, Cordoba, Spain

[3] CRA-W, Walloon Agricultural Research Centre, Gembloux, Belgium

[4] LAGC, Laboratori Agroalimentari de la Generalitat de Catalunya, Barcelona, Spain

[5] NUTRECO, Masterlab Laboratory, Boxmeer, the Netherlands

[6] RIKILT, Institute of Food Safety, Wageningen, the Netherlands

Content:

1. Introduction

In the STRATFEED project, the aim of work package (WP) 5, led by SAC, was to demonstrate the contribution of near-infrared spectroscopy (NIRS) towards controlling compound feed and designing new methods to detect and quantify any addition of animal meal. Six partners (SAC, UCO, LAGC-Gencat, JRC, RIKILT and CRA-W) were involved in carrying out the WP5 activities.

NIRS is probably the most rapid method for testing feed in terms of speed of reporting, timeliness and convenience, allowing a substantial increase in the number of controlled samples and providing an instant response in detecting contaminated specimens. Its speed enables testing to be part of the decision-making processes in managing a feed mill. NIRS is already used widely in the feed industry and is the most likely technique to be used to protect the food chain from contamination by mammalian protein tissue. It is a particularly valuable method for screening feed imported from countries that have not yet experienced bovine spongiform encephalopathy (BSE) and Creutzfeldt-Jakob disease (CJD) and still allow meat-and-bone meal (MBM) to enter the food chain, directly or indirectly (e.g.,

via aquaculture). Our main concern is gross contamination of feed with MBM due to ignorance, malpractice or economic fraud, which can lead to high levels of MBM entering the food chain, thus incurring greater risk. Fish meal (66% CP) is at risk as it commands a very high value, but in the developed world MBM (55% CP) is now a waste product that needs to be disposed. NIRS is important as a method of screening feed to detect suspect specimens and refer them for scrutiny by more labour-intensive forensic tests, such as polymerase chain reaction (PCR), that can be cited in a court of law of an EU Member State.

Although NIRS is easily performed, its calibration and validation present difficulties in that the spectrocomputer must be able to recognise the spectral features of MBM within highly variable plant and fish tissue. As animal fat is still allowed in feed, NIRS must specifically distinguish mammalian protein tissue. This is demanding in terms of the need to include in the model all possible plant-derived ingredients as well as permitted fish meal, fats, oils and dairy by-products. The analytes of importance are those specific to red meat tissue – haemoglobin, myoglobin, collagen, actin and myosin – as well as bone and nervous system tissue.

It must be emphasised that NIRS is a secondary method that depends on a pre-existing primary reference method.

As is well known among NIRS workers, an accurate NIRS measurement of agro-food material relies heavily on the sample set used to develop a mathematical relationship between spectra and the chemical and/or other qualitative reference values (in our case, the percentage of MBM included). The important point about the calibration/training/library set is that samples chosen should represent the full range of characteristics (chemical, physical, technological, etc.). In the STRATFEED project and WP5, great efforts were made to build spectral libraries representing the variability found in European compound feed.

2. Current situation

Several studies have demonstrated the capacity of NIRS technology to detect and quantify animal-origin meal in compound feed. Garrido and Fernández [**5**] reported the results of a tentative evaluation of the potential of the NIRS technology. They used feed samples spiked at various levels (0.5% to 4%) and 'real-process' samples (cattle and pig feed) that also contained 0–4% of MBM, supplied by the feed sector prior to the ban and now no longer available. The authors concluded that NIRS could be used for the instant detection and quantification of MBM at low levels.

Later, Murray *et al*. [**9**] used NIRS to detect MBM in fish meal. PLS discriminant analysis was developed on a calibration set of 67 samples consisting of 22 authenticated fish meal specimens and 45 fish meal specimens deliberately contaminated with MBM at 3%, 6% and 9%. The results showed the potential of this technique to identify animal protein from two species.

Other authors [**4,6,10**] also reported data of calibration equations to quantify the percentage of MBM in compound feed and to classify samples according to the presence/absence of MBM. However, as stated by Garrido and Fernández [**5**], in order to have NIRS models that can be used to label and inspect all feed compounds marketed in Europe, it is necessary to build 'universal' models. The word 'universal' here denotes that the model can be used for a large group of samples and can be applied to any material classed as feed or a feed ingredient fed to farmed livestock. Much of the disappointment that has been expressed about NIRS is due to a lack of

understanding of the 'universal' NIRS model concept. It is also important to clarify here that to build a 'universal' model is not easy; it could take years, and needs European research collaboration, good justification and the support of interested official bodies. This is the type of work carried out in WP5 in the STRATFEED project.

3. Results

The WP5 activities began with designing and writing the protocols for spectra collection, standardisation and transfer by NIRS, undertaken by SAC. These protocols are available on the private STRATFEED website.

3.1 NIRS instrument network

3.1.1 The ISI Standardisation Set Box and the Z score spectra transform

For instrument cloning, a standardisation box consisting of 30 sealed cups of agricultural materials (ISI set) was used (**Table 1**). The box was circulated among the partners, with each one following the protocol designed for checking the instrument performance before conducting duplicate (or more) scans until the spectra were closely matched. Standardisation files to match each instrument were thus constructed. These files would permit spectra from the partners' laboratories to be brought together for developing calibration models.

Table 1: ISI Standard Set (box). Samples in bold indicate animal by-product meal.

1 barley	7 hay	**13 fish meal**	**19 MBM**	25 small grain
2 soy hulls	8 soy meal	14 fresh forage	20 wheat	26 sorghum
3 brewer grain	9 cotton meal	15 gluten feed	21 High protein feed	27 whole soy
4 concentrate	10 distillers	16 gluten meal	22 oats	28 soy meal
5 corn grain	**11 feather meal**	17 hay	**23 poultry BP**	29 Compound feed
6 corn silage	12 mixed feed	18 hay	24 coffee BP	30 wheat

This set of 30 animal feeds and feed ingredients were presented as sealed standard surfaces in all six of the scanning instruments (SAC, CRA-W, LAGC, RIKILT, UCO and NUTRECO), providing the opportunity to survey the large variation among the specimens. The specimens would be expected to show the typical variation encountered universally in animal feed and feed ingredients. The ISI set is thus a suitable model to study. It consists of 26 plant-based feeds and 4 animal by-product meals – feather meal (#11), fish meal (#13), MBM (#19) and poultry meal (#23).

When the spectra of all 30 samples were used to create Principal Component scores (PCs), it was found that a minimum of 8 PCs were required to explain 98% of the variance in the spectra (1100–2500nm; 2,12,2,2 math). These 8 PCs were then used to calculate multivariate 'distances' – Global H (GH) statistics. This was done using either the mean of all 30 specimens as the centre, or MBM, fish meal or poultry meal individually as the centres. The GH distances are shown in **Table 2**.

Table 2: GH distances of the 30 samples in the ISI set. GH values >3 are in red.

Centre:	**MEAN of all ISI 30**		**MBM #19**		**FISH MEAL #13**		**POULTRY MEAL#23**	
position	sample#	GH	sample#	GH	sample#	GH	sample#	GH
1	3	0.103	**19**	0	**13**	0	**23**	0
2	20	0.17	**23**	2.966	**23**	2.315	3	0.739
3	29	0.301	18	**3.155**	3	2.629	29	0.967
4	1	0.314	17	**3.244**	14	2.645	11	1.019
5	12	0.321	20	**3.397**	9	2.649	12	1.149
6	17	0.341	30	**3.42**	20	2.758	20	1.196
7	4	0.381	3	**3.531**	4	2.902	4	1.251
8	7	0.383	9	**3.538**	17	**3.092**	8	1.371
9	6	0.459	4	**3.541**	7	**3.406**	1	1.383
10	8	0.496	7	**3.548**	5	**3.414**	28	1.407
11	28	0.527	26	**3.694**	30	**3.42**	22	1.415
12	9	0.54	15	**3.765**	6	**3.433**	17	1.419
13	5	0.557	1	**3.982**	28	**3.457**	6	1.462
14	18	0.592	25	**4.021**	8	**3.485**	7	1.471
15	15	0.619	6	**4.043**	29	**3.528**	15	1.521
16	22	0.658	8	**4.089**	18	**3.547**	5	1.568
17	**23**	0.703	28	**4.136**	12	**3.554**	24	1.591
18	25	0.718	5	**4.163**	1	**3.596**	9	1.638
19	14	0.973	29	**4.171**	26	**3.635**	18	1.644
20	26	1.119	14	**4.201**	11	**4.032**	10	1.7
21	30	1.142	24	**4.218**	15	**4.074**	25	1.745
22	24	1.183	21	**4.218**	22	**4.103**	14	2.2
23	21	1.205	12	**4.224**	16	**4.147**	**13**	2.315
24	10	1.628	11	**4.51**	25	**4.273**	21	2.407
25	16	1.958	22	**4.736**	27	**4.484**	26	2.416
26	11	2.099	16	**4.943**	10	**4.639**	16	2.467
27	2	2.386	10	**4.976**	21	**4.789**	27	2.602
28	27	2.46	27	**5.428**	2	**4.913**	30	2.694
29	**13**	2.632	2	**6.17**	24	**5.165**	**19**	2.966
30	**19**	**3.03**	**13**	**6.866**	**19**	**6.866**	2	**3.37**

GH calculated from 8 Principal Components; 1100–2500nm; 2,12,2,2 math; no scatter correction

It can be seen that MBM (#19), followed by fish meal (#13), were farthest from the mean of all 30 feeds with GH values, at 3.03 and 2.63, respectively. The GH values can be thought of, somewhat incorrectly, as multivariate versions of a standard deviation. GH values > 3.0 are considered outliers and are in red. When MBM is taken as the centre with a zero value, fish meal is the most remote spectrum (GH=6.87), and poultry meal ((#23) is its closest neighbour (GH=2.97). Nearly all feed specimens have a GH value greater than 3 relative to MBM. When fish meal is taken as the centre with a zero value, MBM is the most remote spectrum (GH=6.87), and poultry meal is its closest neighbour (GH=2.32). It can be concluded, therefore, that both MBM and fish meal are not only markedly different from all other feeds and feed ingredients, but also, and more importantly, very different from one another.

This is because mammalian, avian and fish by-products are rich in lipid, protein and

bone but almost devoid of any carbohydrate, such as starch or cellulose, that constitutes the bulk of plant matter. The reasons why MBM should differ so markedly from fish meal are more difficult to explain. Fish meal is cooked gently at ambient pressure, with the temperature never exceeding 100ºC. MBM is pressure cooked at 133ºC, 3 bars for 20 minutes, which greatly denatures protein and forms Maillard products, reported to cause changes in infrared spectra [**2**]. Because fish meal and MBM are so different, it should be possible to detect MBM in fish meal by infrared spectra, as shown by Murray *et al.* [**9**]. Poultry meal is more problematic because it clusters more closely with the 26 plant-based feeds and is the nearest neighbour to both fish meal and MBM, albeit at a considerable GH distance (2.32 and 2.97, respectively).

An effort was made to identify the underlying spectral differences that may explain why processed animal protein (PAP) differs from plant-derived feed and feed ingredients, and how PAP from different species (mammal, bird and fish) could be distinguished by their reflectance spectra. Identifying the wavelength regions characteristic of different feeds and PAPs is not straightforward because of the ramped baseline offset resulting from scattering. This makes subtraction of spectra or comparison of spectra subjective and prone to error. A new approach was developed during the STRATFEED project. Called Z score spectra, it permits spectra to be transformed and compared.

The mean and standard deviation of the 26 ISI plant-based feeds was used to centre all 30 of the ISI specimen spectra around this mean, so that the spectrum of the mean plant-based feeds lay along a zero line on the wavelength axis. All the individual spectra were then transformed into positive and negative departures from the mean, but scaled in units of the standard deviation. This form of scaling is called 'Z score' – a technique often used to re-scale measurements. Z score is another name for the standard normal variate. It is calculated from:

$$\mathbf{Z}_{\lambda} = (\text{sample}_{\lambda} - \text{mean}_{\lambda}) / \mathbf{SD}_{\lambda}$$

Using this plotting method allowed identification of wavelength bands that were unique to the 4 animal by-product meals and to MBM in particular. While feather meal and fish meal were relatively distinct from MBM, poultry by-product meal was much less well differentiated from MBM. This suggests that poultry meal may be less easily distinguished from MBM by reflectance. This experiment identified wavelengths that could be used to detect MBM (e.g., 1728 nm region of the -CH- stretch first overtone and 2034 nm arising from amide features of denatured protein).

The problem with detecting and quantifying MBM stems from the great diversity of plant and animal material that constitute 'feed' and of the mammalian slaughter by-products known as MBM. This is further complicated by rendered animal fat and fish meal being allowed in feed for certain species only. Most previous analytical experience in NIRS has been in determining the composition of a narrow range of feeds, such as cereal or forage. Both MBM, the analyte to be detected and the feed matrix can be visualised as two diffuse clusters. There are two kinds of detection errors:

- False negatives that risk banned material entering the food chain.
- False positives that incorrectly flag up concern requiring referral / rejection.

At the boundary between uncontaminated and slightly contaminated feed there will be uncertainty and the possibility of not detecting accidental trace contamination. However, blatant fraud contamination usually exceeds 10% inclusion. For this

reason 9% (w/w) was selected as the upper limit for WP5 in artificial mixtures.

Figure 1 shows the Z score spectra of the 30 samples in the ISI set, with the 26 plant-based feeds shown in green. In contrast, the four animal by-product meals are shown in different colours: blue for the fish meal, dark pink for the poultry meal, pale pink for the feather meal, and red for MBM. It is clear that in animal by-product meal there are regions that are distinctly different from plant-based feed. As noted earlier, Z score spectra are scaled in units of ± standard deviation. The Z score spectra for the four types of ISI animal by-product meal are shown in **Figure 2**.

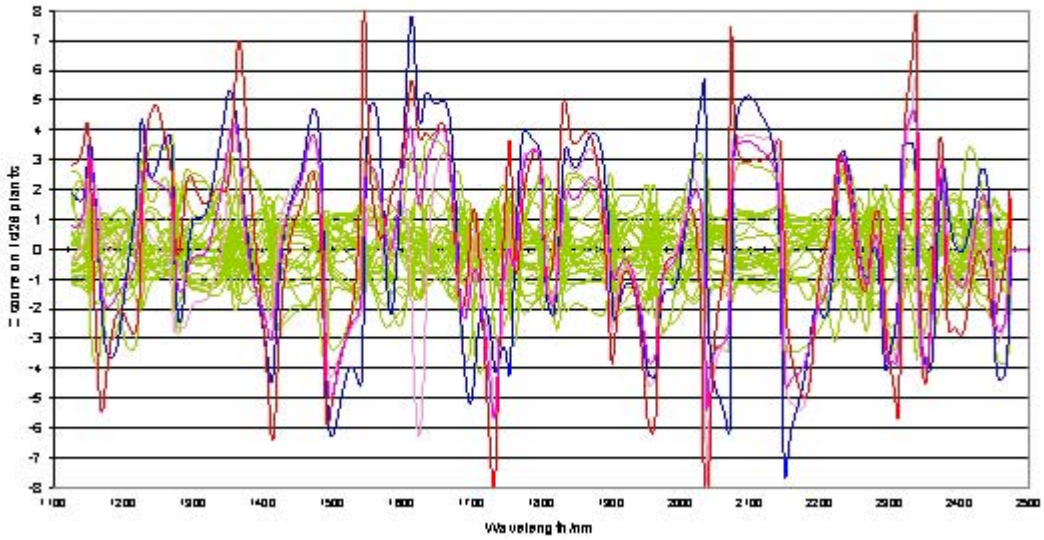

Figure 1: Z score spectra of the 30 ISI standardisation samples. Plant-based feed in green, fish meal in blue, poultry meal in pink, feather meal in pale pink, and MBM in red.

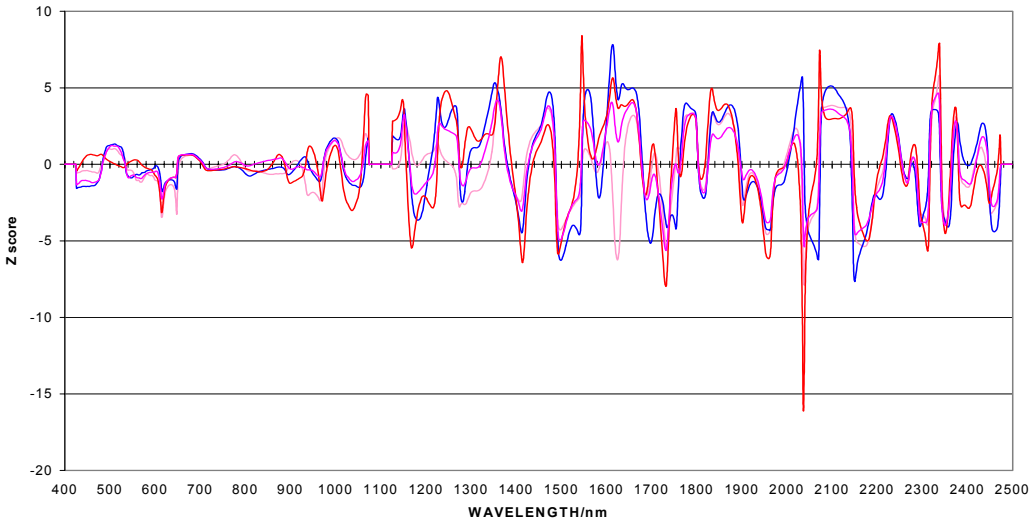

Figure 2: Z score spectra of four types of ISI animal by-product meal, showing unique wavelengths.

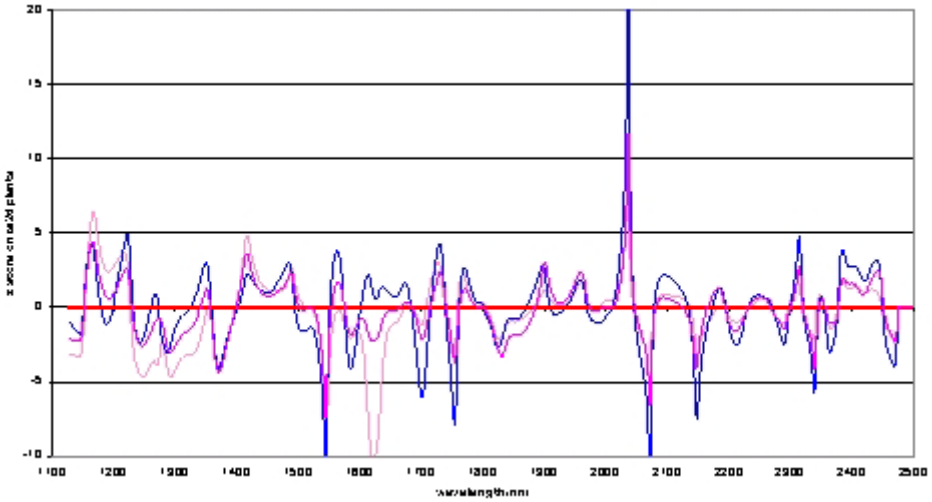

Figure 3: Z score spectra of animal by-products relative to MBM defined as zero.

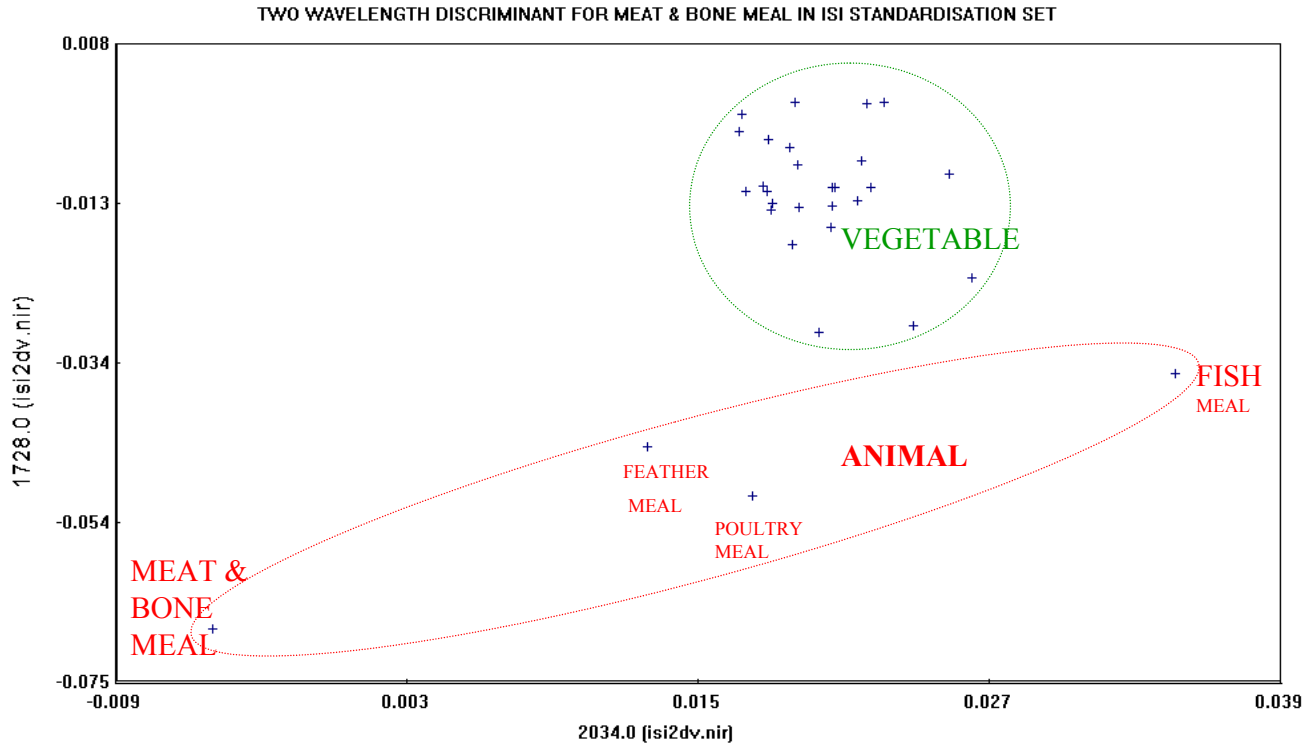

Figure 4: Two wavelength discriminants for animal meal vs vegetable feed.

Figure 3 shows Z score spectra of three types of ISI animal-tissue meal calculated relative to sample 19 (MBM), the Z score being calculated as Z_λ = (sample$_\lambda$ - MBM$_\lambda$)/SD$_\lambda$. The MBM thus lies along the zero line to show wavelengths where MBM clearly differs from other types of animal meal.

Figure 4 shows a two-wavelength (1728 nm vs 2034 nm) plot that best distinguishes fish meal from MBM and from all the other 28 ISI standard specimens. In this plot, MBM and fish meal are widely separated from each other and from the cluster of 26 plant-based feeds. However, the plant-based feed cluster, being very diverse in range, is still quite diffuse,

covering a relatively wide area. This two-dimensional plot shows clearly the problem of detecting MBM in *all* types of feeds. Scaling the levels of MBM present in a feed as a numerical 'distance' score from the centre of this cluster will cause uncertainty because of the unknown position in the cluster of the uncontaminated background feed matrix. This may be described as the 'zero location error'. Thus, unless more additional explanatory wavelengths are able to 'shrink' this cluster, low levels of MBM will be more difficult to detect. Alternatively, if the *all* feeds were restricted to a much narrower subset (e.g., only ruminant feed, or feeds from only one factory), better discrimination could be possible, with less uncertainty at low levels.

This concept explains why macroscopic NIRS is best applied *within* the feed industry in individual factories where an established product spectra library with a timeline, running mean and standard deviation (SD) can act as a basis for detecting suspect specimens among a more *restricted* range of products. In contrast, *micro*scopic NIRS spectra of individual feed particles offers advantages because individual particles of MBM give the spectrum of 100% MBM.

3.1.2 Cloning instruments

A European NIR Spectrometer Network was created to facilitate the merging of spectral libraries of MBM and feed specimens into one large spectral library, in order to harmonised feed evaluation using shared common calibration models.

To transfer the NIRS spectra and equations between the STRATFEED project's NIRS instruments, the standardisation algorithm of Shenk and Westerhaus [**11,13**] available in WinISI software (ver. 1.5) [**7**] was used. The standardisation box described earlier was analysed in duplicate in five instruments at WP5's partner laboratories, following the established protocol. This protocol tries to cover all the factors that could influence this procedure, such as laboratory environment and instrument working conditions. With these data, standardisation matrices (STD files) were developed.

The results of the standardisation were evaluated by comparing the average root mean square of differences corrected for bias (RMS[C]), at *n* wavelengths, between spectra obtained in the master (CRA-W) and satellite instruments before and after standardisation.

$$RMS(C) = 10^6 \times \sqrt{\frac{\sum D^2 - \frac{\sum D^2}{n}}{n-1}}$$

$$D = y_a - y_b$$

where y_a and y_b are the log(1/R) values of two spectra at a given wavelength and *n* is the number of wavelengths used, that is 700 (1100-2500 nm, every 2 nm).

All the spectra scanned on the master instrument (N = 30) were compared with the corresponding spectra scanned on the satellite instruments using the WinISI software Clone program [**7**], following the methods developed by Shenk and Westerhaus [**12, 13**].

From **Table 3** it can be seen that RMS(C) values between spectra obtained after standardisation were similar to those normally encountered by scanning samples in duplicate in one instrument (i.e., 500–700).

To demonstrate the performance of the network, a set of 9 commercial feed samples from set A-21 were analysed in the five NIR instruments described earlier. These samples, provided by the STRATFEED sample bank, were scanned in standard ring cups, apart from the LAGC and SAC instruments where 1/4 rectangular cups and 55 mm-diameter

round cups, respectively, were used. The spectra obtained were predicted using an NIR equation to quantify the percentage of MBM in compound feed by UCO with the set qn-1, to be described later (range: 0.0–34.85%; SD=5.30%; standard error of cross validation [SECV]=0.84%; R^2=0.97). In this case, the UCO instrument was used as master.

Table 3. RMS(C) between spectra before and after standardisation (n = 30), using CRA-W instrument as master.

NIR instrument	RMS(C) before standardisation	RMS(C) after standardisation
SAC	**3854**	**368**
LAGC	**15475**	**1114**
NUTRECO	**9733**	**701**
UCO	**2316**	**503**
RIKILT	**4727**	**359**

Table 4: Statistics for evaluating five NIR instruments, before and after standardisation, using a validation set and an NIR equation for predicting the %MBM added to feed.

	Master	**Satellites**							
	UCO	**CRAW^{before}**	**CRAW^{after}**	**NUT^{before}**	**NUT^{after}**	**LAGC^{before}**	**LAGC^{after}**	**SAC^{before}**	**SAC^{after}**
	2.32	2.52	1.91	-2.46	1.88	3.20	2.59	4.81	3.85
SD	2.61	2.44	2.69	2.43	2.56	2.50	2.68	2.97	3.08
SED(c)	**0.57**	**0.41**	**0.56**	**0.54**	**0.57**	**0.86**	**0.85**	**0.82**	**0.93**
Bias	**-0.27**	**-0.20**	**0.41**	**4.78**	**0.43**	**-0.88**	**-0.27**	**-2.49**	**-1.53**
R^2	0.97	0.98	0.96	0.96	0.95	0.89	0.90	0.93	0.92
Av. H	1.73	3.49*	2.27	6.51	1.89	4.76	1.83	3.31	2.27

[a]SEP(c) value

Table 5: NIRS predicted values (%MBM) of the validation set, before and after standardisation.

Sample	**UCO master**	**CRAW^{before}**	**CRAW^{after}**	**Nut^{before}**	**Nut^{after}**	**LaGC^{before}**	**LaGC^{after}**	**SAC^{before}**	**SAC^{after}**
102	1.3	1.6	0.7	-3.8	0.8	4.0	3.3	4.2	3.2
103	-0.1	-0.4	-1.7	-5.5	-1.5	0.7	-0.6	0.7	-0.5
111	3.1	-	-	-	-	4.6	4.2	4.3	3.3
112	3.5	4.3	3.8	-0.7	3.7	4.7	4.2	7.1	6.3
113	5.2	5.1	4.9	-0.2	4.3	5.8	5.5	8.5	7.7
114	2.9	2.7	2.0	-2.1	2.2	2.8	2.2	4.6	3.6
115	0.4	1.1	0.7	-3.4	0.9	1.4	0.9	3.0	2.1
116	-0.7	-0.3	-0.8	-5.3	-1.1	-0.4	-1.0	1.9	0.8
120	6.3	6.1	5.7	1.1	5.7	6.6	6.1	8.5	7.6

- Error in spectral data

The statistics obtained in the evaluation of the instrument cloning are shown in **Table 4**. Before the standardisation, the bias values ranged from -2.49 to 4.78; after standardisation they fell to between -1.53 and 0.43. The H values ranged from 3.31 to 6.51 before standardisation; after standardisation they ranged from 1.83 to 2.27,showing a convergence in instruments after cloning.Another useful statistic used to evaluate the success of cloning was the H distance [**12**], which is analogous to the Mahalanobis distance. It provides information on the predicted values of the same samples scanned by the satellite and by the master. The results obtained by using the H value confirm those obtained by using the RMS(C) and bias values. The H values obtained in the satellite instruments reach values similar to those found in the master instrument.
The NIRS predicted values of the 9 feeds analysed in the five NIR instruments are shown in **Table 5**.

The results show that, after cloning, the agreement between predicted values in the master and the satellite instruments was improved for each instrument in the network. There was a fall in bias values in all the satellite instruments, apart from the one at the CRA-W laboratory. This shows that the CRA-W spectrophotometer is very similar to the UCO instrument, and that in this case the standardisation had hardly any influence on the results. After standardisation, the reduction in the bias values for the NUT instrument was very marked. Thus, as shown in **Table 5**, before standardisation in the NUT instrument, samples with MBM (i.e., samples 112 and 113) were predicted to be free of MBM. After standardisation, however, the predictions were correct and were very similar to those of the master instrument for the same samples.

In general, the results showed that the STRATFEED NIRS network enables all instruments to produce harmonised results, which is very important for demonstrating that NIRS could be used as a standardised method for implementing the ban on animal meal protein in compound feed throughout Europe.

3.2 Development and validation of NIRS prediction model

3.2.1 Detection and quantification of animal-origin meal in compound feed

The WP5 partners agreed to work with various spectral libraries and to study various strategies of NIR calibration.

The libraries used to develop the prediction models were the ‘real-process’ sample set (set B-NIRS 3) and the ‘experimental’ sample sets (sets B-NIRS 1 and 2 – Centurion C and Trojan T). Set B-NIRS 3 consisted of commercial compound feed supplied by the industry and stored in the STRATFEED sample bank; sets B-NIRS 1 and 2 consisted of samples prepared in the laboratory (preparation details are given in the WP2 part on ‘Sample Bank and Sample Preparation’). Reflectance spectra were acquired and used to form models to detect the presence of mammalian tissue. All the samples were analysed, in ground form, in the spinning and/or transport module in the partners’ various instruments.

The four methods used to detect MBM in feed using the ISI software were:

1. Stepwise Multiple Linear Regression (SMLR), using a few selected wavelengths
2. Modified Partial Least Squares (MPLS) regression, using nearly all wavelengths
3. 'Local', using a database of spectra and selecting in the 'neighbourhood' of an unknown
4. Discriminant, assigning unknowns to two (or more) files ('clean' or ’contaminated’).

All these methods were explored using various mathematical pre-treatments,

particularly the second derivative of log reciprocal reflectance (log1/R). The second derivative is calculated by a running difference between three segments (A, B and C) separated by two gaps. The second derivative is calculated by (A-B)-(B-C) = A - 2B + C. The derivatives most often used in the STRATFEED project were 2,12,2,2 and 2,5,5,1 where the first number was the derivative order (2^{nd} in this case); the second number was the gap in nanometres; and the third and fourth numbers referred to the number of data points in the first and second smoothing. Second derivatives can be easily related to the wavelength space while bringing into focus the local features of a spectrum. Additionally, the software offers six combinations of two types of scatter correction – SNV-DT and MSC– which may be applied before creating the derivative [**1,8**].

Previous UCO research work showed how NIRS could be used to control and detect the illegal addition of animal-origin meal in compound feed [**5,6**]. This work highlighted the importance of calibration set characteristics, such as set size, type of samples ('real-process' versus experimental), the presence of different ingredient matrices and the variability of ingredients used to produce a given formula. These issues could be determining factors for improving the performance of the equations developed.

Therefore, at the start of the project, it was decided that UCO would work with the 'real-process' sample set and that the other partners would work with the experimental sample set (Centurion C and Trojan T). Ultimately, both sets would be merged to develop the 'Global NIR Calibrations'.

3.2.1.1 Calibration using the experimental sample sets (sets B-NIRS 1 and 2 – Centurion C and Trojan T)

3.2.1.1.1 Quantitative strategy

3.2.1.1.1.1 Calibration with set B-NIRS1 - (setC Centurion)

Set C was the first mixing experiment in WP5 of the STRATEFEED project. The procedure for preparing this set is described in the WP2 part on 'Sample Bank and Sample Preparation'.

Set B-NIRS1 produced 900 spectra (triplicate re-packed scans). There were 225 scans at both the 0% and 9% levels, and 105 at each of the 1, 3, 5 and 7% levels, giving a reasonable balance across the concentration range.

Each of the three laboratories involved (SAC, LAGC and CRA-W) scanned samples with three re-packed scan replicates. Spectra were collected by SAC, where clone files were applied to standardise to the CRA-W Foss 6500 monochromator. Both standardised and unstandardised spectra were used in modelling.

Calibration models using MPLS regression on set C usually ran to 16 terms, producing the data shown in **Table 6**. SEC values slightly greater than 1% MBM in feed were obtained for each of the three laboratories. However, when these files were combined, the SEC values almost doubled and the R^2 values fell from 0.9 to 0.7. Standardisation of the files brought no benefit. Loss of performance on combining the partners' files was probably due to increasing diversity among the background feeds accumulated.

Calibrations using 'Local' performed slightly better than MPLS. The combined file did not show the same degraded performance as MPLS because 'Local' does not attempt to produce a model across samples. Rather, it selects a set of nearest neighbours from the database to perform a calibration for each sample likely to be encountered in future.

Table 6a: Results for set B-NIRS1

MPLS 2,12,2,2 (not standardised)

RSQ	SEC	SECV	1-VR	FILE
0.859	1.23	1.297	0.834	**C-SAC**
0.899	1.32	1.474	0.865	**C-CRAW**
0.901	1.00	1.127	0.863	**C-LAG**
0.728	1.93	1.975	0.701	**CENT-ALL**

Table 6b: Results for set B-NIRS1

MPLS 2,12,2,2 (standardised)

RSQ	SEC	SECV	1-VR	FILE
0.859	1.229	1.297	0.834	**#C-SAC**
0.909	1.252	1.431	0.873	**#C-CRAW**
0.901	1.000	1.046	0.901	**#C-LAG**
0.730	1.922	1.940	0.711	**#CENT-ALL**

Table 6c: Results for set B-NIRS1

Local 2,12,2,2 (not standardised)

Npred	Total*	SEP	BIAS	SEP(C)	SLOPE	RSQ	GH	NH	FILE
427	450	1.116	0.03	1.117	1.110	0.884	1.009	0.178	**SAC**
193	270	1.003	0.10	1.000	1.109	0.930	1.082	0.074	**CRAW**
147	180	0.873	-0.11	0.869	1.021	0.912	1.139	0.094	**LAG**
797	900	1.003	-0.01	1.003	1.068	0.919	1.087	0.133	**ALL**

*LOCAL only attempts to predict specimens for which it has enough neighbours

Table 6d: Results for set B-NIRS1

Local 2,12,2,2 (standardised)

Npred	Total*	SEP	BIAS	SEP(C)	SLOPE	RSQ	GH	NH	FILE
427	450	1.094	0.002	1.095	1.113	0.889	1.01	0.178	**#SAC**
245	270	1.006	0.060	1.006	1.063	0.939	1.06	0.069	**#CRAW**
148	180	0.852	-0.09	0.849	1.032	0.915	1.16	0.094	**#LAG**
820	900	1.163	0.012	1.164	1.055	0.895	1.12	0.135	**#ALL**

The oily texture of MBM made successful mixing at best difficult and at worst almost impossible. The partners greatly underestimated the difficulties of this apparently simple preparation task. Containers more than a third full cannot be properly mixed by stirring and/or shaking. In some cases, attempts to ensure good mixing led to agglomeration into 'prill' balls, causing segregation. Subsequent performance of calibration and validation models was almost certainly diminished by inadequately mixed and dispersed oily MBM. Replicated (X3) re-packing and re-scanning did not always show up mixing defects, possibly because the scanned aliquots were too small in mass.

3.2.1.1.1.2 Calibration with set B-NIRS2 - (set T Trojan)

Set B-NIRS2 was the second feed mixing experiment. The aim was to expand the scope of set B-NIRS1 by introducing more 'new' feed samples and to offer sufficient numbers to provide better cross-validation opportunities. Details on set B-NIRS2 are given in the WP2 part 'Sample Bank and Sample Preparation'.

Table 7a: Results for set B-NIRS2

MPLS 2,12,2,2 (not standardised)

RSQ	SEC	SECV	1-VR	FILE
0.585	2.675	2.741	0.529	**T-SAC**
0.872	1.48	1.558	0.848	**T-CRAW**
0.991	0.68	0.716	0.989	**T-LAG**
0.833	2.251	2.284	0.823	**TROJ-ALL**

Table 7b: Results for set B-NIRS2

MPLS 2,12,2,2 (standardised)

RSQ	SEC	SECV	1-VR	FILE
0.582	2.685	2.753	0.525	**#T-SAC**
0.870	1.488	1.565	0.846	**#T-CRAW**
0.991	0.690	0.728	0.989	**#T-LAG**
0.835	2.239	2.292	0.822	**#TROJ-ALL**

Table 7c: Results for set B-NIRS2

LOCAL 2,12,2,2 (not standardised)

Npred	Total	SEP	BIAS	SEP(C)	SLOPE	RSQ	GH	NH	FILE
210	210	1.958	0.015	1.963	1.44	0.838	1.122	0.107	**SAC**
392	420	1.138	-0.08	1.137	1.128	0.929	0.969	0.117	**CRAW**
385	450	0.664	0.033	0.664	1.016	0.991	0.908	0.085	**LAG**
989	1080	1.216	-0.01	1.216	1.06	0.952	0.96	0.097	**ALL**

Table 7d: Results for set B-NIRS2

LOCAL 2,12,2,2 (standardised)

Npred	Total	SEP	BIAS	SEP(C)	SLOPE	RSQ	GH	NH	FILE
209	210	1.945	-0.02	1.95	1.418	0.835	1.12	0.11	**#SAC**
393	417	1.135	-0.09	1.133	1.124	0.929	0.98	0.12	**#CRAW**
386	450	0.701	0.03	0.701	1.018	0.990	0.91	0.08	**#LAG**
969	1080	1.271	-0.02	1.272	1.057	0.947	0.94	0.09	**#ALL**

Calibration models using MPLS on set B-NIRS2 again ran to 16 terms, producing the data in **Table 7**. The performance of the SAC set was very poor (SEC = 2.67%), while the LAGC set performed very well (SEC = 0.68%; RSQ = 0.99). The good

performance of the LAGC set can be attributed partly to the inclusion of specimens with 27% MBM in the model, making the set bimodal, inflating R^2 while reducing SEC to 0.68%. Nevertheless, the LAGC set gave the best calibration statistics, indicating that good mixing and presentation had been achieved. SAC had with mite infestation problems in some samples, requiring heat sterilisation; this affected the results.

Once again, combining three data sets led to degraded performance, but perhaps this can be explained by the poor performance of the SAC set. As in set B-NIRS1, the standardisation of files did not improve performance significantly.

Some models had good calibration statistics, but these tended to decline when more diverse specimens were accumulated. Combining sets from the partners' different laboratory instruments nearly always made performance worse. Calibration models were not robust when cross-validated with other independent sample sets. An alternative approach using the ISI 'Local' procedure was also tried (**Table 7d**). The advantage of 'Local' is that it selects a neighbourhood group of samples surrounding the unknown sample and thus does not depend on fitting all samples into one calibration model. While 'Local' gave good results in calibration, it was not more robust; to be effective, it probably requires a database with a much greater number and diversity of samples. However, it does seem to be a much better procedure for diverse sets of samples than that offered by MPLS regression.

3.2.1.1.1.3 Validation with set B-NIRS5 (setD Decaset)

This was the last laboratory mixture experiment to be performed in WP5. It was conducted to expand the database, to test performance and to provide a basis for validation.

Table 8: NIRS predicted values for Set B-NIRS5

Partner	Code	%MBM-Ref.	%MBM NIRpred.	Partner	Code	%MBM-Ref.	%MBM NIRpred.
SAC	31	**1.00**	**0.12**	Rikilt	11	**9**	**5.16**
SAC	32	**0.00**	**-0.61**	Rikilt	12	**0**	**-3.66**
SAC	33	**3.00**	**1.02**	Rikilt	13	**7**	**3.22**
SAC	34	**0.00**	**0.40**	Rikilt	14	**0**	**-2.41**
SAC	35	**5.00**	**3.22**	Rikilt	15	**5**	**6.79**
SAC	36	**0.00**	**0.57**	Rikilt	16	**0**	**-1.05**
SAC	37	**7.00**	**6.66**	Rikilt	17	**3**	**-1.95**
SAC	38	**0.00**	**4.64**	Rikilt	18	**0**	**-1.64**
SAC	39	**9.00**	**6.42**	Rikilt	19	**1**	**2.34**
SAC	40	**0.00**	**1.93**	Rikilt	20	**0**	**-0.66**
CRAW	13C	**9**	**5.75**	LAGC	01L	**0**	**0.40**
CRAW	18C	**9**	**3.56**	LAGC	02L	**0**	**1.50**
CRAW	21C	**3**	**2.23**	LAGC	03L	**0**	**-0.03**
CRAW	22C	**1**	**1.15**	LAGC	04L	**3**	**3.31**
CRAW	24C	**7**	**1.60**	LAGC	05L	**1**	**-0.48**
CRAW	25C	**5**	**1.26**	LAGC	06L	**9**	**10.52**
CRAW	26C	**3**	**1.98**	LAGC	07L	**0**	**1.72**
CRAW	27C	**1**	**-3.69**	LAGC	08L	**3**	**4.56**
CRAW	29C	**7**	**-10.38**	LAGC	09L	**5**	**4.28**
CRAW	30C	**5**	**2.18**	LAGC	10L	**7**	**5.41**

Validation was carried out using the Decaset prepared in the laboratory and the NIRS prediction model developed with the calibration set made up of sets B-NIRS 1 and B-NIRS 2 (Centurion and Trojan) analysed by LAGC (n=630, SD=6.21, Mean=3.97, SECV=0.84 and R^2=0.98). The results of this in-house validation procedure are shown in **Table 8**. From the 40 samples, six samples were predicted as false positive and three as false negatives. Clearly, poor mixing and presentation played some part in those samples that were incorrectly assigned. Some of these sources of error are explained in a later section.

3.2.1.1.2 Qualitative strategy

Discriminant methods were found to be sufficiently useful for screening purposes, but no one model emerged as reliably robust across the data sets. With the partial least squares (PLS) discriminant, two files ('clean' and 'contaminated') are used to form a model with the file extension .PSD. Unknown specimens are assigned to either file name, depending on which has the larger score. A score of 2.0 is 'perfect identification', while 1.0 is 'no identification'. The sum of both scores is always 3.0, while a score of 1.5 in both file names means that the classification could go either way. An example is shown in

Figure 5 for the PLS discriminant for the SAC set of 450 scans in set B-NIRS1. This model ranged from zero to 9% MBM. The model used 30 factors and gave the correct classification for all specimens. In contrast,

Figure 6 is the combined set B-NIRS1 for all three partners. Of the 900 scans there were 29 false positives and 25 false negatives. Repeated patterns in the scores suggest bias in some background feed samples as the source of these classification defects.

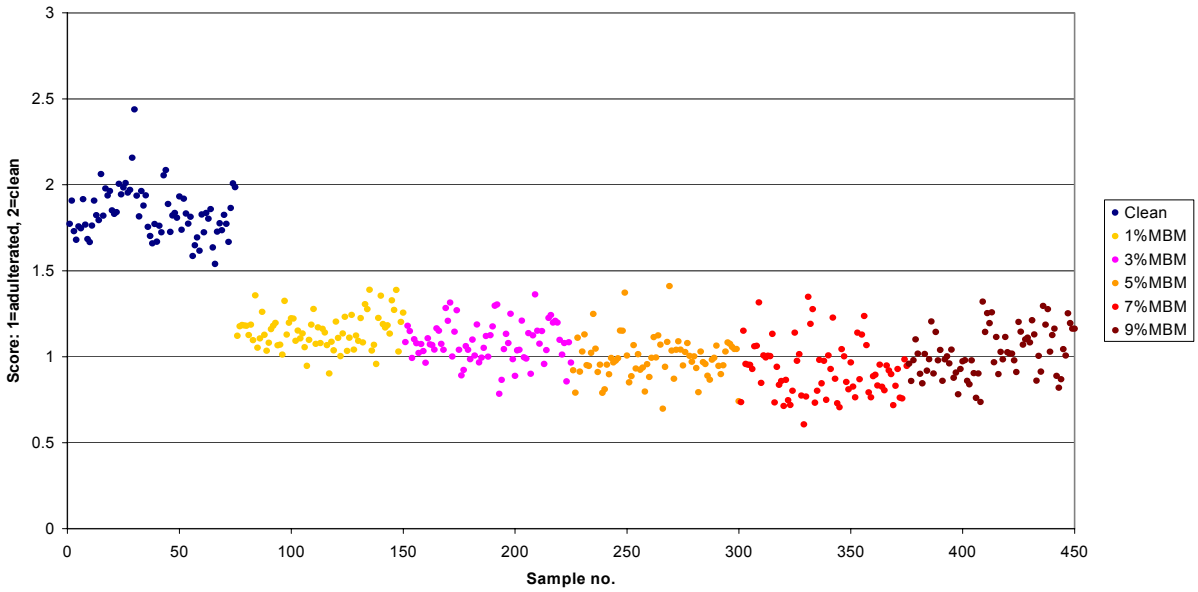

Figure 5: PLS discriminant for %MBM in feed. SAC set B-NIRS1, 30 terms.

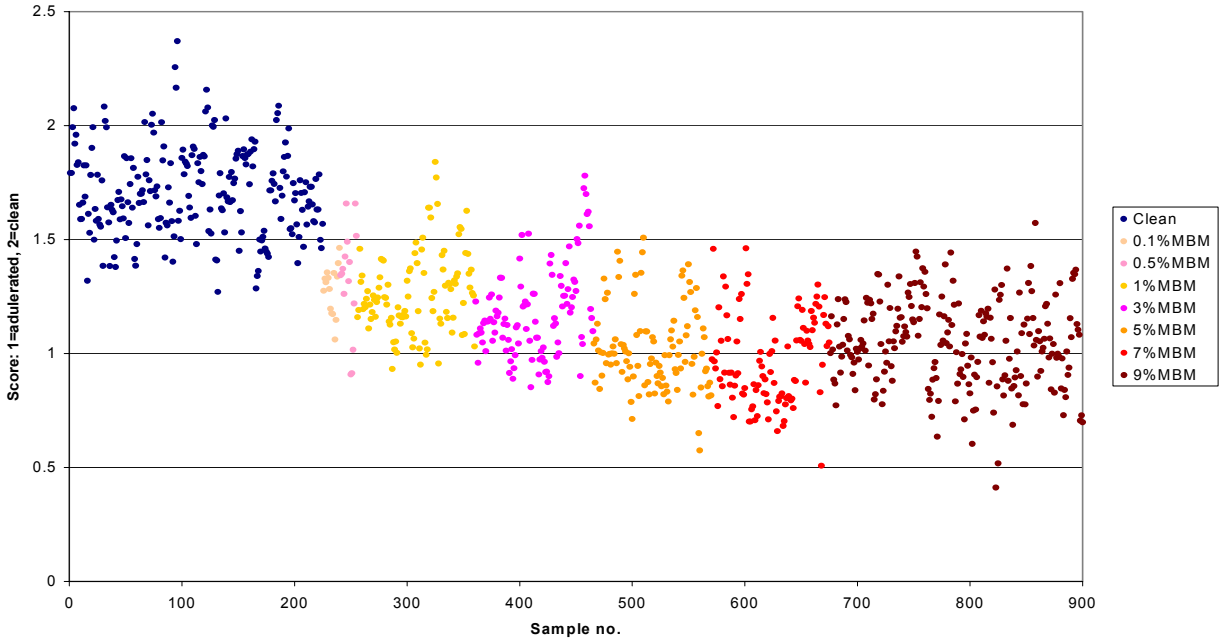

Figure 6: PLS discriminant for %MBM in feed. Combined set B-NIRS1, 27 terms.

Table 9: PLS discriminant results for sets B-NIRS1 et 2, including the ISI box

		False +	False Negative						RSQ
	%MBM:	0	<1	1	3	5	7	9	
Set C	n samples								
CRAW	270	0	0	1	0	0	0	1	0.84
LAGC	180	2	2	0	0	0	0	0	0.84
SAC	450	0	0	0	0	0	0	0	0.802
All	900	29	3	10	10	1	0	1	0.616
Set T									
CRAW	417	2	0	2	0	0	0	1	0.814
LAGC	453	14	2	1	1	0	0	1	0.717
SAC	210	3	0	3	2	0	0	1	0.742
All	1080	84	20	14	21	3	4	8	0.457
Sets C+T									
CRAW	690	5	0	17	3	0	0	1	0.726
LAGC	630	31	5	1	9	1	1	4	0.602
SAC	660	22	0	9	3	1	0	0	0.616
All	1980	258	18	24	15	4	8	9	0.337
C+T+Box									
CRAW	778	13	0	20	4	1	0	1	0.729
LAGC	718	61	18	8	12	0	2	5	0.507
SAC	868	46	0	17	8	4	2	5	0.594
All	2408	242	34	50	28	11	19	47	0.426

The results for the PLS discriminant for sets C and T alone and combined, from all the laboratories, are given in **Table 9**. The number of PLS terms used varied from 18 to 30. Set B-NIRS1 gave the best results, with few errors even when the sets were combined from three laboratories. Combining sets led to poorer performances, in varying degrees. However, those from CRA-W showed the best performance when combined. Of 690 scans, there were only 5 false positives and 21 false negatives, and of 778 scans there were only 13 false positives and 26 false negatives. This reflects the similarity of background feeds used by CRA-W. Combining the ISI set (n=30) did not improve the data in the way we had expected. The results were nevertheless promising in terms of performance.

3.2.1.1.3 Errors arising in mixture experiments and in NIRS calibration

The reliable detection of MBM in feeds depends on finding wavelengths unique to MBM. To be unique, a region must show the largest and most significant difference between MBM and all other acceptable 'clean' feeds. A compromise wavelength region is sought where MBM is most different, while all plant-based feeds are most similar. The Z score transform of spectra is an appropriate discriminant in this context. It maximises the distance between MBM and plant-based feeds while 'shrinking' the cluster of plant-based feeds. Combinations of such regions may perform PLS discrimination. If, however, the background feeds are very diverse, forming a diffuse cluster, this compromise region will lead to a 'zero location error' that affects the low level detection of MBM. It also causes scatter in the predicted values for 'clean' feed, an error that is reflected in those same feeds augmented with MBM.

Another problem is 'distraction'. This occurs because all animal by-products are rich in proteins and lipids. If a model detects proteins and lipids instead of MBM, then all protein- and oil-rich plant foods (e.g., maize gluten feed and oilseed) will show up as suspect. Maillard compounds generated by the high temperatures (133ºC) for rendering MBM may be the source of unique MBM absorption bands arising from denatured protein.

The constant drawback of NIRS is achieving robustness in independent validation sets. Here, calibration (self tests) may appear successful, while validation (blind tests) may not. This arises because regression algorithms tend to 'force-fit' or ‘over-fit’ the data. Overfitting is exacerbated by having too many terms in the model (including spurious correlation) and/or too few samples to adequately represent all possible future samples. Therefore, maximising sample numbers and minimising the number of terms may bring benefits through increased robustness at the expense of some loss in performance.

If proportional mixtures of MBM are carefully prepared for several background plant-feed matrices, regression models can produce a well-fitting model with R^2 approaching 1.0 and SEC approaching zero. However, such models often fail on independent validation; they are not 'robust'. This usually occurs because the independent validation set contains specimens not represented in the calibration set, and/or the number of terms used in the model is too great, leading to overfitting, where spurious correlation occurs. Maximising the number of samples while reducing the number of terms in the model may help through sacrificing good fitting in favour of increased robustness and wider applicability. The ‘Local’ procedure avoids these problems by using a large spectral library that is searched to locate a cluster of spectral nearest neighbours for each specimen tested, including unknown specimens. These nearest neighbours are then used to create a

model to test only that one unknown. This offers distinct advantages. However, the number of library specimens required may be very large before 'Local' is able to test the majority of future feed specimens. If insufficient nearest neighbours are found, 'Local' does not attempt to test the specimen.

Two types of poorly fitted outlier specimens occur: the H statistic and the t statistic. H statistic errors arise because there are 'new' background feeds in the independent validation set. These are spectrally different from those used to construct the model. Such H statistic failures in background feeds will scatter predicted values negatively or positively around zero. This 'zero location error' will persist when the same background feeds are contaminated with low levels of MBM, leading to a scatter of results around the reference value. Poorly mixed or poorly presented contaminated samples will compound this error. In contrast, t statistic outliers arise because of poor mixing and presentation to the instrument (e.g., insufficient area of sample scanned or coarsely ground specimens). Mixing particulate matter with different densities to produce a near homogeneous distribution is difficult to achieve if the particles are very dry or oily and sticky. Sometimes, mechanically mixing or shaking causes segregation. Mixing errors tend to get worse at lower levels of inclusion. It is clear that both outlier types occur in the models created in WP5 of the STRATFEED project.

3.2.1.2 Calibration with the 'real-process' sample set (set B-NIRS-3) and validation with set A-21

UCO worked on improving and optimising chemometric models by testing two complementary calibration strategies: first, the development of qualitative discriminant models for detecting whether or not a feed is contaminated with animal-origin meal; and second, the development of a quantitative model to estimate the percentage of MBM in feeds. All these calibrations were performed using, first, the 'real-process' sample set of compound feed stored in the sample bank and then, in collaboration with other WP5 partners, a 'global calibration file' combining both sets of compound feed, 'real-process' and experimental.

An in-house validation of the models obtained for both strategies was carried out with a blind test set of nine 'real-process' compound feeds, drawn from the STRATFEED sample bank (set A-21). The % MBM in this set ranged from 0% to 6%.

3.2.1.2.1 Qualitative strategy

In the first strategy for developing a qualitative discriminant model, three calibration sets were tested to obtain the prediction equations:

- Set ql-1, with 1,144 samples. The reference data for 560 of these samples were obtained from the formulation declared by the feed company; for the other 584 samples, data were estimated from optical microscopy.
- Set ql-2, with 663 samples. The reference data for this sample set were obtained entirely from the feed formulation.
- Set q1-3, with 1005 samples. This set was a combination of Set q1-2 with the experimental sample set (C+T) scanned on the LAGC instrument and standardised to the UCO instrument.

MPLS regression equations were developed using the three calibration sets described above. A score of 2 was used to identify samples contaminated with animal meal and a score of 1 was used for samples free of animal meal. The results are shown in **Table 10**. The best results were obtained using the calibration set q1-2 in which all the samples came from 'real-process'

commercial feed where the reference data were obtained from the feed formulation. In this set, the SECV was 0.18, while in set ql-1 the SECV value was 0.32 and in set ql-3 it was 0.33. These values correspond to R^2 = 0.87 in set ql-2, compared with 0.57 and 0.54 in the other sets.

*Table 10: Calibration statistics obtained for detecting MBM in compound feed (qualitative strategy)**

	Eqa set ql-1	Eqa set ql-2	Eqa set ql-3
N	1144	663	1005
Mean	1.45	1.54	1.46
Range	1.00–2.00	1.00–2.00	1.00–2.00
SD	0.50	0.50	0.50
SECV	**0.32**	**0.18**	**0.33**
R^2	**0.57**	**0.87**	**0.54**

* These values are discriminant scores, *not* % MBM

The classification of the validation set as contaminated or free of MBM using these qualitative models is shown in **Table 11**. There were two types of classification errors: false positives, where MBM-free specimens are classified as contaminated, and false negatives, where contaminated specimens are classified as free of MBM. As shown, when the model developed with set ql-1 was validated with the 9 blind compound feeds, two mis-classified samples occurred: sample 103 as false positive and sample 114 as false negative. For the model performed with set ql-2, only sample 114 was mis-classified (false negative). The global calibration developed by merging 'real-process' and experimental specimens classified all 9 validation samples correctly; this could be explained by the greater number of feed samples with low MBM levels included in this training set.

Table 11: Detection of MBM contamination in compound feeds for the validation set using the calibrations developed with sets ql-1, ql-2 and ql-3 (qualitative strategy).

Sample	%MBM reference	NIRS predicted values		
		Eqa set ql-1	Eqa set ql-2	Eqa set ql-3
102	**0.9**	+	+	+
103	**0.0**	+	-	-
111	**2.5**	+	+	+
112	**2.5**	+	+	+
113	**4.5**	+	+	+
114	**2.5**	-	-	+
115	**0.0**	-	-	-
116	**0.0**	-	-	-
120	**6.0**	+	+	+

3.2.1.2.2 Quantitative strategy

For the quantitative strategy, to predict the %MBM in compound feed, three calibration sets with different ranges of %MBM were studied.

- Set qn-1, with 630 samples, using the total range of available samples (0.0–34.85%)
- Set qn-2, with 531 samples, using a restricted range (0.0–8.0%)
- Set qn-3, with 1005 samples. This set is a combination of set qn-1 and the experimental sample set (C+T) scanned on the LAGC instrument standardised to the UCO instrument.

The %MBM declared by the feed company was used as reference data. For the experimental samples, the reference data were the %MBM added to the compound feed to produce the mixture. MPLS regression equations were developed using the calibration sets described earlier. Several derivative treatments were tested. In all cases, the NIR spectral range (1100–2500nm) and the SNV and Detrend methods for scatter correction were selected to perform the calibrations. The software WinISI (ver. 1.05) [**7**], was used. The statistics used to select the best equations were the coefficient of determination (R^2), the SECV and the RPD statistic [**14**].

The MPLS calibration statistics for predicting %MBM in compound feed are shown in **Table 12**. The values obtained for SECV, R^2 and RPD confirm the accuracy of the calibrations developed for screening a large collection of samples. However, the predictive performance of the equations may be improved by better coverage of the range with samples of low %MBM. It is therefore clear that, with the available samples, it is more appropriate to use the equation developed with the calibration sets qn-1 or qn-3 (see RPD values), which use the total available range of %MBM.

Table 12: Calibration statistics obtained with three calibration sets for predicting %MBM in compound feed (quantitative strategy).

%MBM	Set qn-1	Set qn-2	Set qn-3
Range	0.00–34.85	0.00–8.00	0.00–34.85
Mean	3.31	1.68	3.24
SD	5.30	2.33	5.50
SECV	0.84	0.72	0.94
R^2	0.97	0.91	0.97
RPD	6.30	3.24	5.85
n	630	531	1005

For the validation set, the values predicted for each model are shown in **Table 13**. The best prediction values were obtained with the calibration set qn-1. As shown, the SEP value obtained with the calibration developed with set qn-1 (0.59%) is half that obtained with set qn-2 (1.18%), and also lower than that obtained with set qn-3 (0.95%).

Most of the samples in the calibration sets used were from Spanish feed companies; the validation samples 102, 103 and 111 were from Belgium. As indicated by the H values (H < 3), all the validation samples

could be predicted without extrapolation of the models, apart from sample 103, which has a rather different spectrum from the others shown by high H values (**Table 13**).

For the calibration developed with sets qn-1 and qn-2, only sample 115 was predicted to be a false positive. The other eight samples were correctly predicted. Nevertheless, with the third equation performed using set qn-3, the %MBM in sample 115 was correct; in this case, only sample 103 was incorrectly predicted. Given that sample 103 had an exceptionally high GH statistic, well above the critical value of 3.0, it is clearly spectrally very different and cannot reliably be predicted by the model. The model did, however, successfully predict the remaining eight samples.

Table 13: Prediction of the validation set using the three calibrations developed (quantitative strategy).

Sample	%MBM Reference	Eqa Set qn-1	Eqa Set qn-2	Eqa Set qn-3	$GH_{set\ qn-1}$	$GH_{set\ qn-2}$	$GH_{set\ qn-3}$
102	**0.9**	**1.3**	**2.3**	**0.6**	1.573	1.464	0.813
103	**0.0**	**-0.4**	**-0.1**	**0.6**	**6.083***	**5.745***	**5.854***
111	**2.5**	**3.1**	**2.0**	**3.8**	1.724	1.462	0.890
112	**2.5**	**3.5**	**3.3**	**3.7**	0.307	0.347	0.229
113	**4.5**	**5.2**	**2.8**	**4.4**	1.280	1.324	1.239
114	**2.5**	**2.9**	**1.9**	**3.5**	0.697	0.750	0.798
115	**0.0**	**0.4**	**0.4**	**-0.1**	0.784	0.722	0.513
116	**0.0**	**-0.7**	**-1.9**	**-0.1**	1.485	1.514	1.161
120	**6.0**	**6.3**	**4.4**	**7.9**	1.629	2.124	1.119
SEP		**0.59**	**1.18**	**0.95**			
Bias		**-0.31**	**0.42**	**-0.60**			

It must be stressed that the models developed with 'real-process' samples should be validated with the same type of samples, because the processing conditions in compound feed manufacture are such that the final product is very different from the artificial mixtures prepared in the laboratory. This was corroborated by other validations of the qualitative and quantitative prediction models carried out with the following sets:

- 54 compound feeds (28 commercial and 26 prepared in laboratory) supplied by the co-ordination team for validation.
- 29 commercial compound feeds supplied by a Spanish feed plant, with the %MBM used in the formulation.
- 10 compound feeds in the SAC 'Decaset', supplied by SAC as blind samples.
- 10 compound feeds, in the RIKILT 'Decaset', supplied by RIKILT as blind samples.

Some of these results were presented at the International Symposium on 'Food and feed safety in the context of prion diseases' held from 16 to 18 June 2004 in Namur, Belgium.

3.2.2 Species identification in animal by-products

UCO has also developed NIRS equations to predict the percentage of MBM derived from tissues of various animal species

(cattle, ruminants, pigs and poultry) in MBM samples.

On 3 October 2002 the EU adopted Regulation EC No. 1774/2002 [**3**], governing Animal By-Products (ABPs). In Article 21 the regulation seeks to address the possible risk inherent in recycling potential infectivity because of the absence of a barrier within species, as well as to exclude cannibalism, which may arise from within-species recycling. Because of that, there is an urgent need to develop methods of analysis which allow identification of the animal species in ABPs and which could overcome the limitations of current analytical methods.

For this application, two strategies were tested: the first was quantitative, relating to predicting the percentage of each animal species in MBM samples, and the second was qualitative, relating to the performance of models for classifying the samples as 'single species specimens' (pig, poultry, cattle) or 'multi-species mixtures'.

The samples used to develop these NIRS prediction models were supplied by a Spanish rendering plant (set B-NIRS4). Each sample was identified according to: sample number; date and time of processing; raw materials used (percentage of tissue of each animal species); sterilisation conditions (temperature, pressure and time); and the name of the person in charge of plant quality control. The calibration set was scanned unground, using the natural product cup, in a Foss NIRSystems 6500 scanning monochromator equipped with a transport module.

For quantitative purposes, MPLS equations were developed to predict the percentages of poultry, pig, cattle, ruminant and non-ruminant in MBM specimens. Various mathematical approaches (derivatives and scatter correction) were used. The best results for each constituent are shown in **Table 14**.

Table 14: Prediction statistics for the percentage of each animal species in MBM.

Constituent	N	Mean	SD	Range	SECV	R^2	RPD
% Poultry meal	204	41.41	35.44	0–100	9.51	0.93	3.73
% Pig meal	189	47.38	31.85	0–100	8.2	0.93	3.88
% Cattle meal	192	9.05	7.75	0–25.63	3.04	0.85	2.55
% Ruminant meal	190	9.37	7.79	0–25.63	2.9	0.86	2.69
% Non-ruminant meal	188	90.85	7.94	74.4–100	3.12	0.85	2.54

For the qualitative approach, a PLS2 discriminant analysis was used. Discriminant models were developed using a training set of 103 samples: 78 of which were in Class I (ruminant, made up of mixtures of different percentages of cattle and other species meal), and the remainder were in Class II (non-ruminant, made up of pure pork and pure poultry meals and one mixture of pork and poultry meal). All the samples in the non-ruminant class in the training and validation sets were correctly classified (**Table 15**), but one sample in the ruminant class was apparently classified incorrectly. However, a subsequent PCR analysis in an external lab (CRA-W) confirmed that this 'mis-classified' sample also contained non-ruminant DNA. The best PLS-DA Type I model was applied to the validation set of 16 samples. **Table 15** shows that all of them were classified correctly.

Table 15: Classification results for species identification (ruminant vs non-ruminant) in MBM.

Training Set Ruminant (n=78); Non-ruminant (n=25)			**Validation Set** Ruminant (n=8); Non-ruminant (n=8)		
Belong to...	Classified as...		Belong to...	Classified as...	
	Ruminant	Non-ruminant		Ruminant	Non-ruminant
Ruminant	**24**	**1**	Ruminant	**8**	**0**
Non-ruminant	**0**	**78**	Non-ruminant	**0**	**8**

These preliminary results show that NIRS technology may allow the animal species present in animal protein by-products to be identified. However, more work is needed to build a sample and spectral data bank with specimens representing the variability found in all types of rendered animal protein meals.
Simultaneously, and linked to this research, UCO is working on the application of NIRS technology for characterising animal fats. Work is in progress to collect authenticated samples to develop robust NIRS prediction models which can authenticate any type of rendered fat produced in Europe.

3.3 Limitations of detection in particulate matter

The limitations of detecting MBM in feed are difficult to determine in solid particulate matter. However, it should be noted that setting unrealistically low levels of detection as a target is misleading. The binomial distribution shows how measurement uncertainty increases as the %MBM decreases. In the following example, 'clean' feed particles are denoted as 'white' and MBM particles are denoted as 'red':

If a feed contains 25% MBM then ¼ of the particles would be red and ¾ would be white. The probability of selecting a red particle is ¼ and of selecting a white one is ¾. If a sample of three particles is selected, the probability that the sample contains 0, 1, 2 or 3 white particles is shown in the expansion below, where the last term represents the all-white-particle ('clean') scenario:

$$(¼ + ¾)^3 = (¼)^3 + 3(¼)^2(¾) + 3(¼)(¾)^2 + (¾)^3$$

$$= \frac{1}{64} + \frac{9}{64} + \frac{27}{64} + \frac{27}{64}$$

So the *expected* frequencies for 64 repeated three-particle samplings are 1, 9, 27 and 27 for 0, 1, 2 and 3 white particles, respectively. Thus, in 64 repeated samplings of 3 particles there will be 27 occasions when MBM would not be detected at all! The last term in the expansion gives the proportion of samplings that would entirely fail to detect MBM.

If the target %MBM is 1%, this last term becomes 99/100 raised to a power of the number of particles examined. If 50 particles were examined, this last term would show that 61% of the sampling tests would show no evidence of MBM in the sample. If the target %MBM is set at 0.1%, then the last term in the expansion (999/1000) would show that 95% of the samplings of 50 particles would show that the feed was 'clean'.

If n feed particles were examined at levels of MBM decreasing from 10% to 0.1%, the percentage of apparently clean samplings would increase, as shown in **Table 16** and **Figure 7**. If 400 particles were examined at 0.1% MBM, two-thirds of the sampling tests would be 'apparently clean'. Thus, as the %MBM decreases, the uncertainty of detection greatly increases for any method that examines particulate matter. Methods that take a large test portion and use density separation for bone tissue are not so prone to this failure, but methods that

take small test portions are much more likely to produce misleading results through unavoidable statistical sampling error.

In a spinning cup presentation, the area scanned is 804 mm^2 corresponding to 1,024 particles with a diameter of 1 mm on the surface layer. In the natural product cell (200 x 16 mm), the area scanned is 3,200 mm^2 corresponding to 4,076 particles with a diameter of 1 mm. These estimates concur with the observed standard errors of estimate being close to 1% MBM for diffuse NIRS.

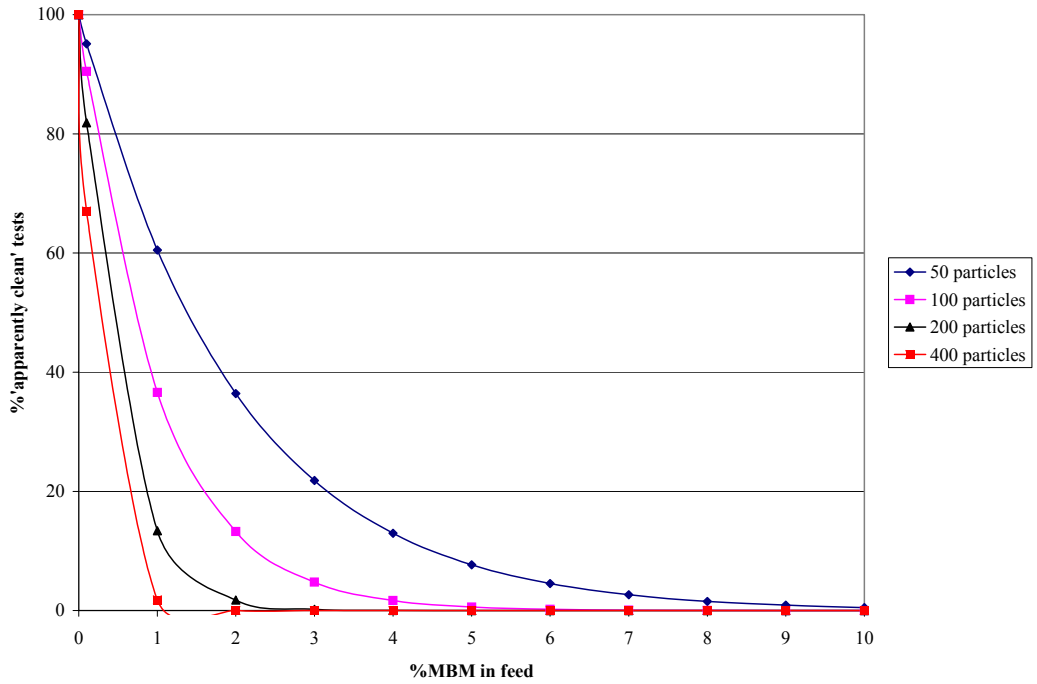

Figure 7: Percentage 'apparently clean' tests of n particles relative to %MBM in feed, calculated from the last term in the binomial expansion.

Table 16: Sampling uncertainty in relation to %MBM particles in feed

%MBM	**Percentage of 'apparently clean ' sampling tests.**					
	50 particles	100 particles	200 particles	400 particles	1,000 particles	4,000 particles
10	0.52	0	0	0	0	0
9	0.9	0.01	0	0	0	0
8	1.55	0.02	0	0	0	0
7	2.66	0.07	0	0	0	0
6	4.53	0.2	0	0	0	0
5	**7.7**	0.59	0	0	0	0
4	**13**	1.69	0	0	0	0
3	**21.8**	4.76	0.03	0	0	0
2	**36.4**	**13.3**	0.23	0.03	0	0
1	**60.5**	**36.6**	1.76	1.8	0	0
0.1	**95.1**	**90.4**	**81.9**	**67**	**36.7**	1.8
0	**100**	**100**	**100**	**100**	**100**	**100**

4. Pros and cons of the NIRS technology

The advantages of NIRS technology are mainly speed, non-destructive analysis, minimal sample preparation, no use of reagents and no production of waste residues, and economical, good and replicable signal intensity coupled with rapid reporting for decision-making processes. The main drawback of NIRS is that it is an indirect method and thus requires large numbers of reference values of authenticated samples to develop prediction models. Also, any future unknown specimen needs to be adequately represented in the calibration model or database, it may not be predicted correctly by the model if its H statistic is beyond the limit (>3.0).

Implementation is beyond the scope of STRATFEED and WP5. However, our research demonstrates that NIRS is best implemented as a screening method in the feed industry, where it is already widely used for routine quality control. Here, there is an established product spectra library with a quality control timeline for a particular product with a running mean and standard deviation. These characterise particular products with a much more restricted variation in composition and spectra, enabling such products to be transformed to a Z score and an H statistic that can detect serious deviations from the average, detect and flag outliers and refer suspect specimens.

In this situation, the main disadvantage of NIRS is resolved at a stroke by assuming that most of the material in circulation is not contaminated. The necessary database is already in the manufacturer's archive. In the pharmaceutical industry such surveillance for inferior drug substitution fraud is commonplace. Legislation should demand that product timeline surveillance becomes a regular part of good manufacturing practice and due diligence. This would benefit responsible manufacturers, their clients and consumers, while lifting the burden on regulatory authorities and allowing them to concentrate surveillance efforts on referred suspect specimens that genuinely warrant scrutiny. Feed is a low-cost commodity with a high volume but a transient existence. Its inspection needs to be appropriate in terms of cost and convenience. If it is not done by NIR, it may not be done at all.

5. Conclusions

The results of WP5 indicate that NIRS could provide the feed industry and inspection bodies with a fast screening method for detecting the contamination of compound feed with animal by-products. NIRS could provide the first line of defence in the food chain and enable more costly methods to be used more productively on suspect specimens. Currently, the detection limit is, at best, 1–1.5%, so it cannot be used alone as legal evidence. Nevertheless, the best results are obtained with the qualitative approach using discriminant analysis rather than with the quantitative strategy. Macroscopic NIRS is best employed in industry where a QC timeline of spectra of a particular product can provide a running mean and standard deviation for a range which is more restricted in terms of composition. In this case, there is a better basis for detecting suspect outlier specimens that warrant further scrutiny.

It has been stressed that it is necessary to build an authenticated sample library that contains samples representing all of the variability in global feed trading. In this sense, also, it is much better to work with ‘real-process’ samples to create robust models that can replicate the type of samples produced by the industry. Therefore, it is necessary to enlarge the spectral libraries constructed in the STRATFEED project with more ‘real-process’ samples to create robust models.

But such samples are now, of course, no longer available due to the ban imposed on processed animal meal. Nevertheless, for this purpose two strategies can be carried out together: first, collecting pet food samples because this type of compound feed is not affected by the MBM ban; and second, producing compound feeds contaminated with MBM in the pilot plants commonly found in large feed companies.
It is also important to note that the results confirm that the cloning procedure enables all instruments to produce harmonised results, demonstrating that NIRS could be used as a standardised method Regarding species identification in animal by-product meal, preliminary results indicate that it is possible to characterise them using NIRS technology. Finally, it should be noted that, on statistical grounds, the setting of unrealistically low detection levels may be unhelpful in practice.

Bibliography

1. Barnes R.J., Dhanoa M.S. and Lister S.J. (1989). Standard normal variate transformation and de-trending of near infrared diffuse reflectance spectra. *Applied Spectroscopy* **43:** 772-777.
2. Cho R.K., Kawano S., Sato T. and Iwamoto M. (1992). Near infrared spectroscopic analysis of impairment of amino acid residues of protein caused by reaction with carbonyl compounds. In: I. Murray and I.A. Cowe (eds) *Making Light Work: Advances in Near Infrared Spectroscopy,* VCH, Weinheim, pp105-111.
3. EC Commission (2002). Regulation No 1774/2002 of 10 October 2002. Official Journal European Communities. 273, 0001.
4. Fernández M., Martínez A., Modroño S. and De la Roza B. (2002). Near infrared reflectance spectroscopy as a tool to predict qualitative and quantitative meat and bone meal presence in compound feed. In: A.M.C. Davies and R.K. Cho (eds.) *Near Infrared Spectroscopy: Proceedings of the 10th International Conference.* NIR Publications, Chichester, West Sussex, UK, pp. 307-312.
5. Garrido A. and. Fernández V. (1998). Report of the workshop on identification of animal ingredients in compound feed focusing on the microscopic method for identification, organised by the CEMA group, 25–26 May 1998. DG XII. Contract SMT4-CT97-6524. Annex X.
6. Garrido A., Pérez-Marín M.D., Guerrero J.E., Gómez A., De Paz F. and Delgado N. (2002). Near infrared reflectance spectroscopy as an essential tool in food safety programmes: predicting ingredients in commercial compound feed. In: A.M.C. Davies and R.K. Cho R.K. (eds.) *Near Infrared Spectroscopy: Proceedings of the 10th International Conference*. NIR Publications, Chichester, West Sussex, UK, pp. 145-150.
7. ISI (2000). The complete software solution using a single screen for routine analysis, robust calibrations, and networking. Manual. FOSS NIRSystems/TECATOR. Infrasoft International, LLC. Silver Spring, MD, USA.
8. Martens H., Jensen S.A. and Geladi P. (1983). Multivariate linearity transformations for near-infrared reflectance spectroscopy, In: O.H.J. Christie (ed). *Proceedings of the Nordic Symposium on Applied Statistics.* Stokkland Forlag: Stavanger, pp. 205-234.
9. Murray I., Aucott L.S. and Pike I.H. (2001). Use of discriminant analysis on visible and near infrared reflectance spectra to detect adulteration of fishmeal with meat and bone meal. *J. Near Infrared Spectrosc*. **9**, 297-311.

10. Piraux F. and Dardenne P. (1999) Feed authentication by near-infrared microscopy. Eds Giangiacomo, R. and al. *Proceedings of the 9th International Conference on Near-Infrared Spectroscopy*. Verona, It, June 1999, pp 535-541
11. Shenk J.S.(1992). Networking and calibration transfer. In: I. Murray and I.A. Cowe (eds) *Making Light Work: Advances in Near Infrared Spectroscopy*. VCH, Weinheim, Germany, pp. 223-228.
12. Shenk J.S. and Westerhaus M.O. (1995). Routine operation, calibration development and network system management manual. NIRSystems, Inc., 12101 Tech Road, Silver Spring. MD 20904. USA.
13. Shenk J.S. (1990). Standardizing NIRS instruments. In: R. Biston and N. Bartiaux-Thill (eds) *Proceedings of the 3rd International Conference on Near Infrared Spectroscopy* Agricultural Research Centre Publishing, Gembloux, Belgium, pp. 649-654.
14. Williams P.C. and Sobering D. (1996). How do we do it: a brief summary of the methods we use in developing near infrared calibrations. In: A.M.C. Davies and P.C. Williams (eds.) *Near Infrared Spectroscopy: The Future Waves*. NIR Publications, Chichester, West Sussex, UK, pp. 185-188.