Quantitative Characterization of Feed and Agricultural Samples Using a Handheld Near-Infrared Analyzer

S.K. Schreyer, Lin Zhang, Thermo Fisher Scientific, Tewksbury, MA, United States

Background

Near-infrared (NIR) spectroscopy has been widely accepted for use in the food and agricultural areas, beginning with the work of Karl Norris at the USDA to develop quality methods for agricultural products. However, since NIR spectra are overtones and combinations of fundamental IR spectra, the peaks tend to lose definition and broaden, representing general features due to CH, NH and OH stretch and bend frequencies, as contrasted to their fundamental frequencies that define IR. Thus, it was not until mathematical and computer modeling of NIR spectroscopy was utilized, that information could be obtained from NIR. With the implementation of chemometrics methods, a valuable tool for differentiation and quantification of agricultural and food components became available. Coupled with the reduced requirement for sample preparation, one of the major strengths of NIR and the implementation of NIR in a portable, handheld unit such as the Thermo Scientific™ microPHAZIR™ analyzer, NIR spectroscopy and the subsequent identification and quantification of food, feed and agricultural samples can be taken from the laboratory directly to the field or warehouse.

NIR spectroscopy is well recognized as a reliable instrument to predict moisture, protein and fat in food or agricultural samples. One of the most common uses – dating from the earliest implementation of NIR spectroscopy – is the quantification of protein, moisture and ash in flour[1,2]. We present results obtained from the quantitative study of quality parameters in feed and ingredients based on the Aunir INGOT™[3] library. Prediction models for quality parameters were developed from the INGOT library by Aunir. Further work was performed in-house to develop algorithms for robust predictions including bias correction, model optimization and calibration transfer between instruments. Further, as these models will be deployed on several handheld microPHAZIR analyzers, it is important to determine the effect of model transfer amongst the same type of instruments. This work builds upon a recent study published by Dardenne[4] to evaluate the potential of a calibration transfer from the Foss NIRSystem 6500 to the handheld microPHAZIR analyzer, and our internal efforts to develop robust calibration models for feed and feed ingredients.

The collaboration with Aunir allows us to utilize the well-established and robust INGOT library that was developed over several years. By utilizing our handheld portable NIR instruments the predictive models can be moved from the requirement for a lab-based environment into the field, thus increasingly opening the availability of this widely recognized and useful method.

Study Scope

In the following discussion we will present the progress of this study from the preliminary assessment of a small subset of the library to the resulting predictive models. Initially models were built in-house to test a subset of wheat and soya samples, and to develop the proof of concept. Based on the success of these models for prediction of common physical parameters, a collaborative study was done with Aunir, using their Ingot library and the microPHAZIR analyzer. (Appendix A INGOT Library) In this latter study, Aunir supplied 14 models for feed and feed ingredients based on the Ingot calibration library. In collaboration with Aunir these models were validated, refined and using in-house methods, calibration transfer and bias correction were performed. In the following sections some of these methods and the results will be presented.

Instrumentation

All testing was done using the portable microPHAZIR AG analyzers, and with a specially designed sample cup for Food, Feed, Agriculture (FFA) application. The sample cup can be attached to the instrument and manually rotated to multiple positions. The sampling window on the instrument is designed to locate at an off-the-center position such that each rotation of sample cup results in a different portion of sample being presented to the instrument. The samples were scooped and sampled a given replicate times and the predicted results were averaged as indicated.

Thermo SCIENTIFIC

Materials

All materials tested on the microPHAZIR AG analyzer units were used as received from Aunir. These were ground samples that covered the range of properties appropriate to the model. These materials were used to test the capability of the microPHAZIR analyzer for quantitative analysis of FFA parameters. In total 11 parameters are the most common constituents of interest in FFA applications and they are listed in Table #1.

Predicted Parameter	Description	
Moisture		
Oil A	Fat (EE)	ether extract
Oil B	Fat (AH)	acid hydrolysis
Protein		
Fibre	crude fibre	
Ash		inorganic matter
Starch		enzymatic starch
Sugar	Reducing sugar	
NCGD	neutral cellulose plus gamanase	
NDF	neutral detergent fibre	hemicellulose + ADF
ADF	acid detergent fibre	cellulose, lignin, fibre-bound N

Table 1 Most commonly predicted parameters for FFA applications

For actual application, not every parameter could be predicted for each model. Model parameters are restricted to the most useful and practical for quality evaluation relevant to certain ingredient or feed sample type. Of particular relevance to most applications is the prediction of protein, moisture and fat (oils). In the following sections the focus will be on prediction of the most common parameters relevant to the material being tested.

Results and Discussion

Part 1. Sample preparation and spectra collection

In the initial testing phase the applicability of the unit to predict moisture, protein and fat was evaluated using an initial set of materials received from Aunir. This set consisted of 20 samples each of ground wheat, soya and corn, covering an appropriate range of parameters.

The parameters evaluated for wheat and soya are indicated in Table #2 and Table #3. For initial testing wheat was used as an indication of the performance of cereals (Aunir Group 10) and soya was used as indication of the performance of High Protein, Low Oil (Aunir Group 30)

All samples were used as received (ground) and placed in the quartz sample cup for NIR spectra collection. Spectra were collected on microPHAZIR AG analyzers, each over a wavelength range from 1595-2395 nm, in diffuse reflectance mode. Spectra were collected over six positions of the sample cup in order to compensate for sample inhomogeneity. In total, this resulted in 6 spectra collected per sample, and each sample was also tested 3 times, with replacement, resulting in 24 spectra collected for each sample. This sampling process was repeated for each of the 20 samples. Samples were scanned in a randomized manner to compensate for any sampling correlations. All sampling was completed the same day to deter any day to day variations.

The spectral data were then evaluated and quantitative individual PLS-1 models were constructed using our internal chemometrics software package Thermo Method Generator™ software (TMG). This software was developed for use with the microPHAZIR analyzer.

An example of the spectra collected on each microPHAZIR analyzer is shown in Figure 1.

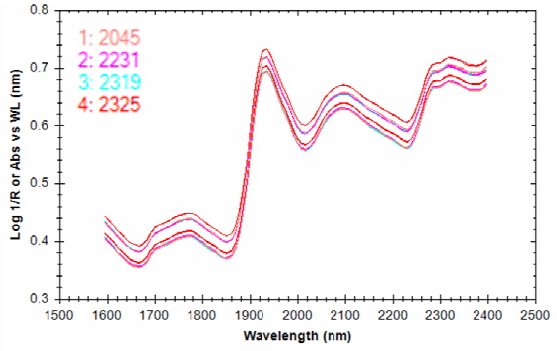

Figure 1 Example of wheat spectra from 4 different instruments

Wheat	Moisture	Oil A	Oil B	Protein	Starch	Sugar
Mean	13.14	1.29	2.32	12.69	57.82	3.49
Range	2.12	0.51	0.77	5.90	8.59	2.96
Minimum	12.39	0.94	1.77	10.01	54.48	1.52
Maximum	14.51	1.45	2.54	15.91	63.07	4.48

Table 2 Reference values for wheat

Soya	Moisture	Oil A	Oil B	Protein	Starch	Sugar
Mean	11.70	1.67	2.37	49.84	4.84	10.01
Range	2.44	1.89	1.83	6.19	1.83	4.50
Minimum	10.44	0.70	1.46	45.69	4.03	7.95
Maximum	12.88	2.59	3.29	51.88	5.86	12.45

Table 3 Reference values for soya

Aside from baseline offset, all spectral features were similar across the different microPHAZIR analyzers, with no obvious spectral non-conformities. Further results will be presented using one analyzer.

Based on one microPHAZIR analyzer, the resulting spectra collected from the 20 wheat samples are shown in Figure 2.

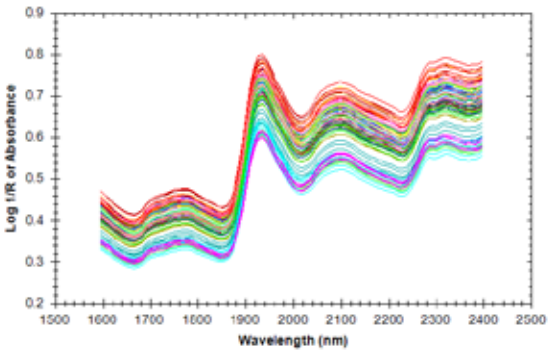

Figure 2 Spectra of wheat across the full range for protein reference values

Part 2. Model development example (Protein)

Preprocessing was performed using standard normal variate (SNV) to offset particle inhomogeneity and particle density and packing differences, followed by Savitsky-Golay smoothing (1st derivative, 7 point smooth, 2nd order polynomial). The effective wavelength region used for protein determination was adjusted in each case. For protein, the wavelength was restricted to 1716.7- 2359.6 nm. These regions include the N-H overtone and combination bands. Results are on Figure 3.

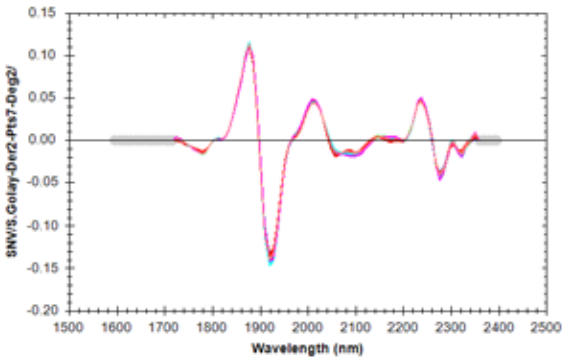

Figure 3 Preprocessed spectra of wheat samples

It was determined that 3 factors were optimal for the PLS model, based on the plot of factors and associated root mean square error (RMSE) of cross validation. Associated loadings plots for the first 3 factors also substantiate the use of 3 factors in the resulting PLS model as past 3 factors loadings plots show increased loss of information and increased noise. Factors indicate the importance of the CH combination bands at 1700, the overtone bands at 2200-2300nm and the nitrogen overtone and combination bands at 2000-2200 regions.

The resulting PLS model gave a RMSE of calibration of 0.25% and a RMSECV of 0.27% for protein prediction. The predicted results gave a R^2 of 0.97, as shown in the correlation plot for prediction of protein across the 20 calibration samples on Figure 4.

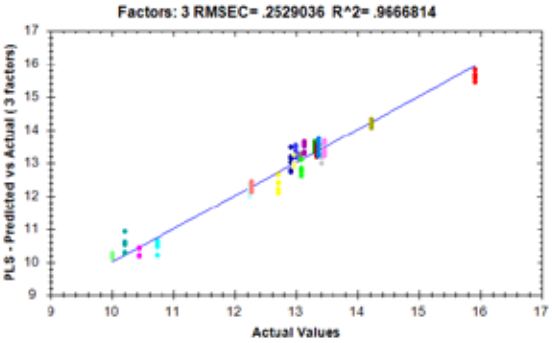

Figure 4 Correlation plot of the reference and the predicted values for wheat protein

Further refinement of the model can be made by omitting the water peak and restricting the wavelength regions to 1716.7-1900,2000- 2359.6nm. However the improvement in RMSE was relatively small and had little effect on the overall model, so this is not necessary unless it appears that the model would be required to function in varying moisture sample environments. For a proof of concept model we did not include this wavelength restriction as all testing was done in the lab.

A residual plot of the predicted results for both the calibration data set shows the variation associated with protein levels at each sample collection. Results are presented on Figure 5. This is associated with the sample inhomogeneity and is expected for this type of sample material. It merely highlights the importance of taking this variation into account for sample collection, by the use of prediction averaging. In sample averaging, results will be reported as an average of a given set of scans. This will be presented in the following section.

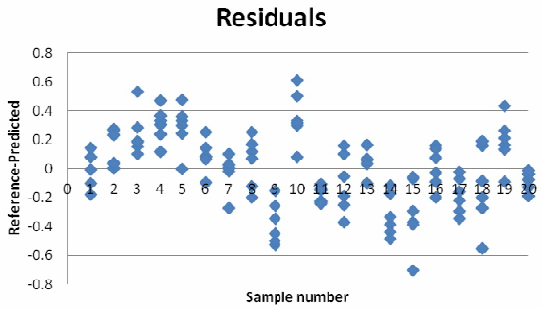

Figure 5 Residuals (actual-predicted) for the wheat samples

Part 3. Proof of concept: Predicting wheat and soya parameters

Similar models were built for the other parameters for wheat (moisture, oil, sugar and starch). Each individual model was evaluated for the optimized preprocessing conditions and once determined, each model was then evaluated for the calibration accuracy. Based on this, models were built for prediction of each parameter. These individual models were then combined together into a master multi-PLS model to be used for wheat prediction.

The preceding results and discussion was based on the calibration data set. In order to ascertain the functionality of the model built, and assess the predictive ability, the wheat application was then loaded onto four independent microPHAZIR AG analyzers.

	Predicted results for Wheat Sample				
Wheat	Moisture	Protein	Oil A	Sugar	Starch
Reference	12.97	13.08	1.44	3.52	56.88
FFA 2319	12.97	12.70	1.35	3.70	58.71
FFA 2231	13.02	12.51	1.47	4.05	58.40
FFA 2045	12.81	13.04	1.32	3.72	58.87

Table 4 Predictions for wheat samples from 4 instruments with comparison to the reference values

	Predicted results for Wheat Sample				
Wheat	Moisture	Protein	Oil A	Sugar	Starch
Reference	12.97	13.08	1.44	3.52	56.88
FFA 2319	12.97	12.70	1.35	3.70	58.71
FFA 2231	13.02	12.51	1.47	4.05	58.40
FFA 2045	12.81	13.04	1.32	3.72	58.87

Table 5 Predicted results for the soya samples across 4 instruments with comparison to the reference values

Predictions were obtained for a series of runs. Samples were scooped into the sample cup and for each sample 6 predictions were made; each prediction after a rotation of the sample cup. This was repeated with the same sample on all 4 units. Then new samples were used and the above procedure was repeated two more times. This gave a total of 18 predictions per unit, and an overall number of 54 predictions on protein. The averaged results for prediction of associated parameters in wheat are shown below for each of the 4 microPHAZIR analyzer units in Table 4.

A similar procedure was followed for soya analysis. The individual models were made based on data collected from the microPHAZIR analyzer units, and evaluated individually to optimize model accuracy. These individual models were then combined into a soya multi-PLS model, and loaded onto the 4 units. In the case of soya, predictions were made for moisture, protein, both types of oil, sugar and starch. Predictions were generated similar to the wheat models – predictions were averaged over 6 positions and three samples – and compared to the reference values supplied by Aunir. Prediction results are shown for soya in the table in Table 5.

Part 4. Evaluation of Aunir models: Cereals

The first library and model evaluated from Aunir was Group 10 (Cereals). This comprises data from 8 cereal types including wheat and corn. This model was built by Aunir using their internal INGOT library (see Appendix for library list) of 30,000+ samples collected on traditional benchtop NIR instrument with high spectral resolution and larger spectral range than microPHAZIR analyzer. To augment the library with microPHAZIR analyzer data, additional samples were collected using two parent microPHAZIR analyzer instruments at Aunir. The library from benchtop NIR was transferred to match the microPHAZIR analyzer platform and transferred spectra were combined with the data collected on the parent units. From this data set, models were developed for prediction of quality parameters on cereal samples. 20 wheat samples were measured on four instruments. For each sample, a single compaction of sample is loaded into sample cup with three evenly spaced rotations during sampling.

Tables 6, 7 and 8 summarize the prediction performance of Group 10 model on 20 wheat samples using four instruments. Three metrics are presented in the tables, i.e., SEP, bias and bias corrected error. These metrics and their

Sampling Pattern	Constituent	Unit 2319			Unit 2325			Unit 2395			Unit 2398		
		SEP	Bias	Bias Corrected Error	SEP	Bias	Bias Corrected Error	SEP	Bias	Bias Corrected Error	SEP	Bias	Bias Corrected Error
Averaging	Moisture	0.59	-0.46	0.37	0.60	-0.44	0.41	0.43	-0.06	0.43	0.46	-0.07	0.46
	Oil A	0.29	-0.24	0.17	0.57	-0.55	0.14	0.52	-0.50	0.16	0.53	-0.51	0.16
	Oil B	0.41	-0.38	0.15	0.67	-0.66	0.13	0.68	-0.66	0.15	0.62	-0.60	0.16
	Protein	0.31	0.12	0.28	0.29	-0.05	0.29	0.30	0.01	0.30	0.39	0.16	0.36
	Fiber	0.95	-0.86	0.40	1.57	-1.51	0.42	1.24	-1.14	0.48	1.28	-1.20	0.45
	Ash	0.28	-0.06	0.28	0.31	-0.18	0.25	0.25	-0.02	0.25	0.29	-0.13	0.25
	Starch	1.86	0.76	1.70	3.25	2.72	1.78	2.05	0.61	1.95	2.44	1.54	1.90
No Averaging	Moisture	0.60	-0.46	0.39	0.61	-0.44	0.43	0.45	-0.06	0.45	0.49	-0.07	0.48
	Oil A	0.31	-0.24	0.20	0.58	-0.55	0.19	0.54	-0.50	0.20	0.54	-0.51	0.20
	Oil B	0.42	-0.38	0.17	0.68	-0.66	0.18	0.69	-0.66	0.18	0.63	-0.60	0.19
	Protein	0.36	0.12	0.34	0.32	-0.05	0.32	0.38	0.01	0.38	0.45	0.16	0.42
	Fiber	0.97	-0.86	0.44	1.58	-1.51	0.46	1.25	-1.14	0.52	1.29	-1.20	0.48
	Ash	0.30	-0.06	0.29	0.32	-0.18	0.27	0.26	-0.02	0.26	0.30	-0.13	0.27
	Starch	1.98	0.76	1.83	3.31	2.72	1.89	2.15	0.61	2.06	2.54	1.54	2.02

Table 6 Model prediction performance from 4 instruments on the 20 wheat samples to compare averaging vs not averaging

Sampling Pattern	Constituent	Average across 4 units		
		SEP	Absolute Value of Bias	Bias Corrected Error
Averaging	Moisture	0.53	0.26	0.42
	Oil A	0.49	0.45	0.16
	Oil B	0.60	0.58	0.15
	Protein	0.33	0.08	0.31
	Fiber	1.28	1.18	0.44
	Ash	0.28	0.10	0.26
	Starch	2.46	1.41	1.84
No Averaging	Moisture	0.54	0.26	0.44
	Oil A	0.50	0.45	0.20
	Oil B	0.61	0.58	0.18
	Protein	0.38	0.08	0.37
	Fiber	1.29	1.18	0.48
	Ash	0.30	0.10	0.27
	Starch	2.55	1.41	1.95

Table 7 Model prediction performance based on average of the 4 instruments. SEP and bias results included to compare averaging across samples and not averaging

performance in this application are explained below.

SEP is the standard error of prediction and is an assessment of overall error in prediction. SEP includes both systematic error and random error. Bias is calculated as the mean difference between model predicted and reference values and is an estimate of systematic error in prediction. For the FFA application, bias could be caused by several factors. First, there might be a difference in wet chemistry test methods for reference values. For example, a customer may prefer a certain type of protein analysis which might have a constant bias relative to the wet chemistry used by INGOT library. This difference is not addressed in the study here but could be easily removed by the on-board bias and slope correction software. The on-board software can be configured to apply customer bias and slope to the predicted concentration of a constituent of interest. Second, there is always some difference between parent and child instrument. This difference could be a result of the different way light propagates from the instrument and the sample, irrespective of how tight manufacturing control is. This difference results in bias of the predicted parameter. For quantitative applications of NIR, some NIR instrument manufacturers perform some kind of instrument standardization to improve prediction accuracy. Our results below show that bias across instruments are negligible in the context of FFA application. Thus, no instrument standardization is performed for the current release. Third, the calibration model is built upon a group of similar samples (barley, corn, wheat, rye etc.) and here the prediction is performed on a specific sample type (wheat). In theory, the model presents no bias when an imaginary "averaged" sample from different sample groups is predicted. However, a small systematic bias is expected when one specific sample type is predicted. Again, this could be addressed by on-board bias/slope correction if needed. Bias corrected error is calculated by removing the contribution of bias from SEP. This metrics represents the error cause by uncertainty in measurement system itself. Thus, proper sampling could further reduce this error.

The performance of a model across four instruments was shown to be satisfactory. For example, in the case of protein, the SEP values range from 0.29 to 0.39 with bias ranging from -0.05 to 0.16. The mean SEP is 0.33 with a mean of absolute bias of 0.08. The bias corrected mean SEP is 0.31.

With regard to sampling error, comparing the case of averaging versus non-averaging (averaging over multiple positions over sample cup), the bias corrected error is reduced 0.37 to 0.31. Since this error is part of SEP, the corresponding reduction in SEP is 0.05.

	max bias	min bias	Range of Bias Among instruments
Moisture	-0.06	-0.46	0.40
Oil A	-0.24	-0.55	0.31
Oil B	-0.38	-0.66	0.28
Protein	0.16	-0.05	0.21
Fiber	-0.86	-1.51	0.65
Ash	-0.02	-0.18	0.16
Starch	2.72	0.61	2.11
Moisture	-0.06	-0.46	0.40
Oil A	-0.24	-0.55	0.31
Oil B	-0.38	-0.66	0.28
Protein	0.16	-0.05	0.21
Fiber	-0.86	-1.51	0.65
Ash	-0.02	-0.18	0.16
Starch	2.72	0.61	2.11

Table 8 Comparison of bias between the 4 instruments for wheat sample prediction

Summary

The use of a portable, handheld NIR instrument to predict protein, moisture, fat and other parameters on feed and agricultural ingredients has been shown to give reliable and robust results. Results were shown as a progression from proof of concept through to final optimized models using protein and moisture analysis in some detail. In either case, a robust and useable prediction model was achieved, with relatively low prediction errors.

The models from Aunir are very robust and no calibration standardization was needed. Some instrument bias was observed but it is expected that the on-board bias/offset correction software could be used to fine-tune the predictions. Further the sampling error could be reduced by firstly grinding sample before scanning over multiple positions.

References

[1]Osborne B.G., Fearn T., 1983 Journal of Food Technology, 18, 453-460

[2]Norris, K.H., Williams, P.C., 1984 Cereal Chemistry 61(2), 158-165

[3]Aunir INGOT Animal Feed Library, Towcester, UK; info@aunir.co.uk

[4]Daredenne, P., Pierna J. A. F., Vermeulen, P., Lecler, B., Baeten, V., 2010 Applied Spectroscopy, 64(6), 644-648

Level 3	Feed					
Level 2	**Poultry** Group 1	**Ruminant** Group 2	**Swine** Group 3	**Equine** Group 4	**Concentrates** Group 5	**Aqua Feed** Group 6
Level 1	Broiler	Beef	Finisher Pig	Racehorse	Poultry Concentrate	Fresh Water
	Chick	Sheep	Gestation Sow	Pony	Broiler Concentrate	Salt Water
	Duck	Calf	Grower's Pig	Horse	Duck Concentrate	
	Game	Dairy	Piglet		Layer Concentrate	
	Goose	Dairy Blends			Ruminant Concentrate	
	Layers	Lamb			Diary Concentrate	
	Ostrich	Goat			Blend Concentrate	
	Turkey				Pig Finisher Concentrate	
					Pig Grower Concentrate	

May 2013

Level 3	Ingredients								
Level 2	**Cereals** Group 10	**High Protein High Oil** Group 20	**Hi Protein Low Oil** Group 30	**Legumes** Group 40	**Low Protein High Oil** Group 50	**Low Protein Low Oil** Group 60		**Animal Protein** Group 70	**Milk Powders** Group 80
Level 1	Barley	Rape-Pulse Mix	Corn Gluten Feed 60%	Beans	Biscuit Meal High Oil	Citrus	Palm Kernal Low Oil	Fish Meal	Dried Whey Powder
	Corn (Maize)	Rape Seed Expeller	Cotton Extract	Peas	Maize Germ Meal	Cocoa	Potato Products	Meat and Bone Meal	Skimmed Milk Powder
	Oats	Soya Full Fat	Groundnut Extract	Lupins	Palm Kernal High Oil	Copra	Rice	Blood and Plasma	Full Cream Milk Powder
	Triticale	Sunflower Seeds	Linseed Extract	Alfalfa Lucerne	Rice Bran High Oil	Copra Extract	Rice Bran Extract	Feather Meal	
	Wheat	Cotton Seed Expeller	Malt Residue			Corn Gluten Feed 20%	Grain Screenings	Poultry By-Products	
	Red Wheat	Linseed Expeller	Rape Meal Extract			Distillery Low Protein	Sugar Beet	Bone Meal	
	Sorghum	Whole Rape	Soya Meal Extract			Locust Bean	Wheatfield Middling	Crab	
	Rye	Sesame Expeller	Sunflower Extract			Soya Hulls	Barley Bran	Shrimp	
		Soybean Expeller	Sesame Meal Extract			Oat Feed	Maize Bran		
		Whole Soybean	Distillery High Protein			Oat Flour	Biscuit Meal Low Oil		
		Whole Linseed				Grass Meal	Maize Germ Meal		
						Hominy	Shea Nut Meal		
						Cassava			

May 2013

ISO 9001 REGISTERED COMPANY
Thermo Fisher Scientific, San Jose, CA USA is ISO Certified.

thermoscientific.com

© 2013 Thermo Fisher Scientific Inc. All rights reserved. INGOT is a registered trademark of Aunir.
All other trademarks are the property of Thermo Fisher Scientific Inc. and its subsidiaries.
Specifications, terms and pricing are subject to change. Not all products are available in all countries. Please consult your local sales representative for details.

Africa-Other +27 11 570 1840
Australia +61 2 8844 9500
Austria +43 1 333 50 34 0
Belgium +32 53 73 42 41
Canada +1 800 530 8447
China +86 10 8419 3588
Denmark +45 70 23 62 60
Europe-Other +43 1 333 50 34 0
Finland /Norway/Sweden +46 8 556 468 00
France +33 1 60 92 48 00
Germany +49 6103 408 1014
India +91 22 6742 9434
Italy +39 02 950 591
Japan +81 45 453 9100
Latin America +1 608 276 5659
Middle East +43 1 333 50 34 0
Netherlands +31 76 579 55 55
South Africa +27 11 570 1840
Spain +34 914 845 965
Switzerland +41 61 716 77 00
UK +44 1442 233555
USA +1 800 532 4752

Thermo SCIENTIFIC
Part of Thermo Fisher Scientific

AN12345_E 03/13W