



## XVII<sup>th</sup> World Congress of the International Commission of Agricultural and Biosystems Engineering (CIGR)

Hosted by the Canadian Society for Bioengineering (CSBE/SCGAB)  
Québec City, Canada June 13-17, 2010



### FAST DETECTION OF CRUDE PROTEIN CONTENT IN FISH FEEDS BASED ON INFRARED TRANSMISSION SPECTROSCOPY AND CHEMOMETRICS

FENGLE ZHU<sup>1</sup>, PENGCHENG NIE<sup>1</sup>, YONG HE<sup>1</sup>

<sup>1</sup> F. Zhu, College of Biosystems Engineering and Food Science, Zhejiang University 268 Kaixuan Road, Hangzhou, 310029, China, fengle7777777@yahoo.cn

<sup>1</sup> P. Nie, niepengc@163.com

<sup>1</sup> Y. He, yhe@zju.edu.cn

#### CSBE101447 – Presented at Section VI: Postharvest Technology, Food and Process Engineering Conference

**ABSTRACT** Infrared spectroscopy and chemometric methods were investigated to determine crude protein content in fish feeds fast and non-destructively. 15 different brands were prepared for infrared transmission spectroscopy measurement using a FT/IR-4100 Fourier Transform Infrared Spectrometer (350-7800 cm<sup>-1</sup>). A total of 150 samples were randomly selected to create the calibration model, and the remaining 75, with similar distribution of crude protein content, were used to verify the model. Multivariate analysis techniques including partial least squares (PLS) and least square-support vector machine (LS-SVM) were developed to construct the crude protein component calibration models simultaneously. Grid research and leave one out cross validation procedures were employed for optimization of LS-SVM parameters. Compared with PLS, an eximious prediction precision was achieved based on LS-SVM with correlation coefficient (R) = 0.9955 and root mean square error for prediction (RMSEP) = 0.8918 based on mid infrared spectral data. Furthermore, the comparison of prediction results showed the performance of models with mid infrared (MIR; 400-3999 cm<sup>-1</sup>) spectral data was better than that with near infrared (NIR; 4000-6666 cm<sup>-1</sup>) spectral data. Based on this study, near/mid infrared transmission spectroscopy combined with LS-SVM could be promising to be applied as a simple and alternative way for the fast quantitative assessment of crude protein content in fish feeds.

**Keywords:** Near/mid infrared spectroscopy, Fish feeds, Crude protein, Least squares-support vector machine, Partial least squares

**INTRODUCTION** Fish meat is popular nutritious food and part of economic products which earns a large amount of exporting incomes to China. In fish farming, the feeds costs account for 40-60% of the total investment and have been manufactured in an increasing quantity year by year in China. Fish feeds which take a significant part in fish growth, metabolism and reproduction, are leadingly consisted of protein, carbohydrate, lipid and moisture. At present, a myriad of fish feeds are available on the market, also prevalent are some inferior and adulterated fish feeds which contain inadequate nutritional content. Thus the quality control should be executed to standardize the feeds quality by the concerned authority to crack down on counterfeit goods and protect the interests of consumers and legal manufacturers.

As crude protein is the most significant element for quality control in fish feeds, its content is a critical issue for feeds suppliers, feeds researchers, fish farmers and resource consultants. The currently employed methods for protein determination are mainly chemical analyses, such as Kjeldahl method (Helrich, 1990) and Dumas combustion method (Etheridge et al., 1998). Although the attained results are precise and universally approved, the standard procedures are usually complicated, expensive, polluttional, time-consuming and demanding professional staffs. Hence, a keen demand exists for seeking an instant and accurate quantificational technique for the crude protein content detection in fish feeds to substitute for the conventional methods.

Infrared spectroscopy, such as near infrared (NIR) spectroscopy and mid infrared (MIR) spectroscopy, is regarded as a rapid and dependable way of investigating the quality and safety of agro-food products (Downey, 1998). NIR spectroscopy is an analytical technique on organic compound, which reflects the componential and constructional information of samples at spectral range of 4000-12500  $\text{cm}^{-1}$ . The analytical abilities of NIR depend on the repetitive and broad absorption of the near infrared light by C-H, O-H, N-H, C-O and C-C bonds (van Raamsdonk et al., 2007). Nowadays, with the advantages of rapidity, accuracy, high efficiency, excellent repeatability, simultaneous measurements, environmental friendliness, limited sample preparation, absence of solvents usage and reduced long-term maintenance costs for qualitative and quantitative analyses, NIR-based methods are acquiring extensive acceptance in many industries, such as food, agriculture, pharmaceuticals, cosmetics, textiles, petroleum and chemical industry (Yan et al., 2005).

Lots of papers have demonstrated the ability of the NIR technology to predict the crude protein values for different animal feeds and forages efficiently. De Boever et al. (1995) applied NIR to predict protein content in cattle feeds based on partial least square (PLS) effectively. Murray et al. (2001) and Yang et al. (2008) demonstrated the potential of the NIR technique for the rapid and reliable detection of processed animal protein in fish feeds. NIR spectroscopy models were developed on the basis of PLS by Xiccato et al. (2003) to predict the crude protein in rabbit feeds. The crude protein values in compound feedingstuffs for different animals were determined using NIR calibrations by Pérez-Marín et al. (2004). Maneerot et al. (2006) performed multiple linear regression (MLR) and PLS to assess crude protein in laboratory prepared shrimp feeds using NIR. González-Martíns et al. (2006) estimated the crude protein content in various animal feeds instantaneously and non-destructively using modified PLS based on NIR technology and a remote reflectance fiber optic probe.

MIR spectroscopy could be utilized to acquire compositional discrepancies among samples and proffer information regarding the fine chemical structure of samples based on fundamental vibrations of various chemical groups at spectral range of 400-4000  $\text{cm}^{-1}$  (Reid et al., 2005). Potential analysis capability of MIR spectroscopy in food science has been certified (Downey et al., 1998; Irudayaraj et al., 2003; Flåtten et al., 2005). Nonetheless, so far there has been little published information relating to the application of MIR spectroscopy to assess chemical contents in fish feeds.

Both NIR and MIR techniques, nevertheless, have several limitations. On the one hand, spectral bands in the NIR region are overtone and combination bands deriving from fundamental bands found in the MIR region, leading more noises, overlaps and unrelated information to NIR, rendering it more difficult to interpret (Reid et al., 2005). Besides, as the NIR bands are more severely superimposed than MIR bands, the molar absorptivity of MIR bands is higher than NIR bands, and the peaks of MIR

are sharper, more specific, and more sensitive (Chung et al., 1999; Shao et al., 2008). Thereby, with more information of intensities and frequencies which are richer and stronger than NIR, MIR can monitor lower levels of key analytes and functional groups (Roychoudhury et al., 2006). On the other hand, compared with MIR, the signal-to-noise ratio, combined with the spectral reproducibility of NIR is superior (Chung et al., 1999). Moreover, NIR with fiber optic probe for diffuse reflectance measurements could be executed with little sample preparation and controlled remotely which make the whole operation more convenient (Wang et al., 2006). Therefore, the richness and usefulness of spectral information, signal-to-noise ratio, reproducibility of measurement and easiness of operation should be considered synchronously for appropriate selection of a spectroscopic mean for analyses.

The major objectives of the present study were to (i) appraise the feasibility of determining crude protein in fish feeds using infrared spectroscopy, (ii) contrast the prediction precision between least square-support vector machine (LS-SVM) and PLS analyses, (iii) compare the performance of NIR and MIR spectroscopy based upon chemometrics.

## MATERIALS AND METHODS

**Materials** 15 brands of fish feeds with a broad range of crude protein content were purchased from the local market. All samples were randomly divided into calibration set of 150 samples and prediction set of 75 samples. No single sample was used in calibration and prediction sets in the meanwhile. Samples in the two sets were kept unchanged to compare the performance of different calibration models. After grinding and drying, the reference method for determining crude protein content was conducted by Dumas combustion method (Sweeney, 1989) and the factor 6.25 was used to convert the nitrogen value to crude protein content. About 100mg samples were weighed into a porcelain sample holder for introduction into the combustion chamber utilizing an automated sample loader.  $N_2$  was converted from covalently bound nitrogen through the combustion process, and quantitated through the gas passing by a conductivity cell (Etheridge et al., 1998). The detailed information of crude protein content in samples was revealed in Table 1.

**Spectral Acquisition and Preprocessing** A JASCO model FT/IR-4100 Fourier Transform Infrared Spectrometer (Jasco, Tokyo, Japan) with a valid range of 350-7800  $cm^{-1}$  and resolution of 4  $cm^{-1}$  was applied for transmittance rate measurements. The ambient temperature was kept at  $20 \pm 1$  °C constantly during the experiment. After grinding to pass through a 100 mesh sieve, samples were mixed with KBr at the ratio of 1:49 with a mortar and pestle for the sake of enhancing the transmission rate. Prior to measuring by the spectrometer with an average of 64 successive scans for each spectrum, the mixture was compressed into a uniform tablet with thickness of 2mm and diameter of 5mm. On account of system imperfections and external disturbances, apparent scattering noises were observed at the beginning and end of the spectral curves. To improve the measurement accuracy and signal-to-noise ratio, the MIR and NIR analyses were based on 400-3999  $cm^{-1}$  and 4000-6666  $cm^{-1}$  respectively. The software Spectra Manager CFR was used for spectral measurement and analysis.

Before calibration, the transmittance spectra were transformed into absorbance data through  $\log(1/T)$  algorithm, where T= transmittance. Then MIR and NIR spectral data were preprocessed by multiplicative scatter correction (MSC) (Isaksson and Næs, 1988) and standard normal variate (SNV) (Barnes et al., 1989), respectively, implementing by The Unscrambler 9.7 (CAMO PROCESS AS, OSLO, Norway).

**Least Squares-Support Vector Machine** LS-SVM which is an optimized algorithm based on the standard SVM regression proposed by Suykens et al. (2002), could handle with linear and nonlinear multivariate analyses and resolve these problems in a relatively fast way (Suykens and Vandewalle, 1999). It applies a set of linear equations instead of a quadratic programming (QP) problem to acquire the support vectors (SVs). Thereby, LS-SVR not only has the merit of good generalization performance like SVM but also possesses a simpler structure and shorter optimization time. When adopting LS-SVM, three crucial problems were required to figure out, namely, the determination of optimal input data set, adequate kernel function, and best LS-SVM parameters. The whole spectral data with pretreatment were used as input data set in this study. For selection of kernel function, no systematic methodology was available. Radial basis function (RBF) kernel was recommended as the kernel function of LS-SVM due to the fact that it was a nonlinear function and a more compact supported kernel, and capable of decreasing the computational complicity of the training procedure while giving good performance under general smoothness assumptions. The setting of optimized parameters, including regularization parameter  $\gamma$  and the RBF kernel function parameter  $\sigma^2$  played a valuable role in setting up a good LS-SVM regression model.  $\gamma$  which was significant in improving the generalization performance of LS-SVM model, determined the tradeoff between structural risk minimization (SRM) principle and empirical risk minimization (ERM) principle.  $\sigma^2$  as the bandwidth of RBF kernel function was related to the confidence of the data, and reflected the distribution of  $X$  values of training data (Cherkassky et al., 2004). Its adjustment also influenced the nonlinear nature of the regression, that is, small values of  $\sigma^2$  yielded a more nonlinear model, which might result in overmany regressors and over-fitting; however, with the increasing of  $\sigma^2$ , the kernel became wider, forcing the model toward a more linear solution (Chauchard et al., 2004) which might lead to reduced regressors, simpler model and more inaccurate results. Details of the LS-SVM description could be found in literature (Guo et al., 2006). The free LS-SVM toolbox (LS-SVM v 1.5, Suykens, Leuven, Belgium) was applied with MATLAB 7.0 (The Math Works, Natick, USA) to derive the calibration models.

**Partial Least Squares** PLS is a comprehensively utilized multi-analysis and bilinear modeling method in spectroscopic analysis (Geladi et al., 1986). It is a procedure applied to modeling the relationship between the spectral variables matrix ( $X$ ) and the target chemical properties matrix ( $Y$ ) for predicting with the smallest number of LVs. The  $Y$  information was used during the decomposition process so that the  $X$  containing higher constituent concentrations could be weighted more heavily than those with lower concentrations (Wu et al., 2009). It is particularly suitable with more variables than samples. PLS was executed using the Unscrambler V9.7, with full cross validation to validate the quality and prevent over-fitting of the calibration models.

**Model Evaluation** The prediction performance for all developed models was chiefly evaluated by correlation coefficient ( $R$ ) between the measured and predicted values, root mean square error of prediction (RMSEP) (Galvão et al., 2008), and bias. Generally, an ideal model should have a higher  $R$  value and lower RMSEP and bias values.

## RESULTS AND DISCUSSION

**Overview of the Spectra** Throughout the NIR measurement area (4000-6666  $\text{cm}^{-1}$ ), the trends of different curves were relatively parallel and no obvious peaks and valleys existed. However, spectral transmission curves were more complex in MIR region (400-3999  $\text{cm}^{-1}$ ), which were displayed in Figure 1. The amide I vibration,

absorbing near  $1650\text{ cm}^{-1}$ , arose mainly from the C=O stretching vibration with minor contributions from the out-of-phase C-N stretching vibration, the C-C-N deformation and the N-H in-plane bend. The amide II band around  $1550\text{ cm}^{-1}$  was the out-of-phase combination of the N-H in-plane bend and the C-N stretching vibration with smaller contributions from the C-O in-plane bend and the C-C and N-C stretching vibrations. The amide III band between  $1400\text{-}1200\text{ cm}^{-1}$  was the in-phase combination of the N-H bending and the C-N stretching vibration with small contributions from the C-O in-plane bending and the C-C stretching vibration (Barth, 2007). The phosphate groups covalently bound to casein proteins around  $1060\text{-}1100\text{ cm}^{-1}$ . Moreover, the obvious troughs around  $2925$ ,  $2854$ ,  $1747$  and  $1160\text{ cm}^{-1}$  embodied characteristic valleys of fat. Area between  $1250$  and  $800\text{ cm}^{-1}$  which represented various C-O vibrations in carbohydrates, mainly lactose (Zhou et al., 2006) was not apparent. Areas around  $1640$  and  $750\text{ cm}^{-1}$  stood for two large water bands,  $\nu_2$  and  $\nu_L$  (Shao et al., 2008). However, troughs around  $1650\text{ cm}^{-1}$  and  $1640\text{ cm}^{-1}$  were overlapped. In both NIR and MIR regions, the spectral curves did not exhibit an obvious rule, and multivariate calibration procedures were required to extract useful information related to spectral data and analyze the properties of fish feeds.

**Comparing Prediction Results of LS-SVM and PLS Models** LS-SVM was performed with RBF kernel for crude protein evaluation. For the selection of the optimum combination of  $(\gamma, \sigma^2)$ , a two-step grid-search technique and leave one out cross validation were employed to avoid over-fitting. Grid-search was a two-dimensional minimization procedure based on exhaustive search across the specified range using geometric steps. The first step of grid search was for a crude search with a comparatively large step in a  $10 \times 10$  grid. Subsequently, the second step was for the specified search with a smaller step in a  $10 \times 10$  grid to obtain the optimal pair of  $(\gamma, \sigma^2)$ . The optimization ranges of  $\gamma$  and  $\sigma^2$  were both  $10^{-4}\text{-}10^4$  for MIR data,  $10^{-4}\text{-}10^6$  for NIR data. For leave one out cross validation, data were once permuted randomly. In every iteration, one left one point, and fitted a model on the other data points. The model performance was evaluated based on the point left out. For each data point this procedure was repeated, and all the different estimates of the performance were combined eventually. PLS models were established simultaneously with the same sample sets. The prediction performances of LS-SVM models outperformed PLS models in both MIR and NIR ranges (Table 2).

**Comparing Prediction Results of MIR and NIR Bands** The prediction results based on different spectral ranges were shown in Table 2, which indicated that both LS-SVM and PLS models built with MIR data turned out to be much better than NIR data. The original spectral transmission plots showed that the transmission rate curves of NIR range were extremely similar and the main difference was offset. Besides lots of high frequency noises and low sensitivity of the signal in NIR region, another disadvantage was the superposition of many different overtone and combination bands which caused lower structural selectivity for NIR spectra. In MIR region, curves were smoother and more specific with many obvious peaks and valleys, and there were great differences among different brands. The primary cause for these better relationships in MIR than NIR region was the stronger absorptions of crude protein in MIR which were associated with fundamental valence vibrations of functional groups of the molecule (Karoui et al., 2005). The relatively high concentration of crude protein in fish feeds was another contributing factor.

**CONCLUSION** Quantification of crude protein in fish feeds was achieved using infrared transmission spectroscopy combined with chemometric methods. Regression

models produced more acceptable precision in prediction based on MIR than NIR. Further investigations should be done to enlarge the number of brands and samples, as well as improve the model's robustness and rigorousness before its online application.

**ACKNOWLEDGEMENTS** This work was supported by the 863 National High Technology Research and Development Program of China (2007AA10A210), Zhejiang Provincial Natural Science Foundation of China (Z3090295), and Agricultural Science and Technology Achievements Transformation Fund Programs (2009GB23600517).

## REFERENCES

- Barnes, R. J., M. S. Dhanoa, and S. J. Lister. 1989. Standard normal variate transformation and de-trending of near infrared diffuse reflectance spectra. *Appl. Spectrosc.* 43: 772-777.
- Barth, A.. 2007. Infrared spectroscopy of proteins. *Biochim. Biophys. Acta* 1767: 1073-1101.
- Chauchard, F., R. Cogdill, S. Roussel, J. M. Roger, and V. Bellon-Maurel. 2004. Application of LS-SVM to non-linear phenomena in NIR spectroscopy: development of a robust and portable sensor for acidity prediction in grapes. *Chemometrics Intell. Lab. Syst.* 71: 141-150.
- Cherkassky, V., and Y. Q. Ma. 2004. Practical selection of SVM parameters and noise estimation for SVM regression. *Neural Netw.* 17: 113-126.
- Chung H., M. S. Ku, and J. S. Lee. 1999. Comparison of near-infrared and mid-infrared spectroscopy for the determination of distillation property of kerosene. *Vib. Spectrosc.* 20: 155-163.
- de Boever, J. L., B. G. Cottyn, J. M. Vanacker, and Ch. V. Boucqué. 1995. The use of NIRS to predict the chemical composition and the energy value of compound feeds for cattle. *Anim. Feed Sci. Technol.* 51: 243-253.
- Downey, G. 1998. Food and food ingredient authentication by mid-infrared spectroscopy and chemometrics. *Trends Anal. Chem.* 17: 418-424.
- Etheridge, R. D., G. M. Pesti, and E. H. Foster. 1998. A comparison of nitrogen values obtained utilizing the Kjeldahl nitrogen and Dumas combustion methodologies (Leco CNS 2000) on samples typical of an animal nutrition analytical laboratory. *Anim. Feed Sci. Technol.* 73: 21-28.
- Flåtten, A., E. A. Bryhni, A. Kohler, B. Egelanddal, and T. Isaksson. 2005. Determination of C22:5 and C22:6 marine fatty acids in pork fat with fourier transform mid-infrared spectroscopy. *Meat Sci.* 69: 433-440.
- Galvão, R. K. H., M. C. U. Araújo, W. D. Fragoso, E. C. Silva, G. E. José, S. F. C. Soares, and H. M. Paiva. 2008. A variable elimination method to improve the parsimony of MLR models using the successive projections algorithm. *Chemometrics Intell. Lab. Syst.* 92: 83-91.
- Geladi, P., and B. R. Kowalski. 1986. Partial least-squares regression: a tutorial. *Anal. Chim. Acta* 185: 1-17.
- González-Martín, I., N. Álvarez-García, and J. L. Hernández-Andaluz. 2006. Instantaneous determination of crude proteins, fat and fibre in animal feeds using near infrared reflectance spectroscopy technology and a remote reflectance fibre-optic probe. *Anim. Feed Sci. Technol.* 128: 165-171.
- Guo, H., H. P. Liu, and L. Wang. 2006. Method for selecting parameters of least squares support vector machines and application. *J. Syst. Simul.* 18: 2033-2036-2051.
- Helrich, K.. 1990. Official methods of analysis of the association of official analytical chemists. 15th ed. Arlington, America: Association of Official Analytical Chemists

- Inc.
- Irudayaraj, J., F. Xu, and J. Tewari. 2003. Rapid determination of invert cane sugar adulteration in honey using FTIR spectroscopy and multivariate analysis. *J. Food Sci.* 68: 2040-2045.
- Isaksson, T., and T. Næs, 1988. The effect of multiplicative scatter correction (MSC) and linearity improvement in NIR spectroscopy. *Appl. Spectrosc.* 42: 1273-1284.
- Karoui, R., É. Dufour, L. Pillonel, E. Schaller, D. Picque, T. Cattenoz, and J. O. Bosset. 2005. The potential of combined infrared and fluorescence spectroscopies as a method of determination of the geographic origin of emmental cheeses. *Int. Dairy J.* 15: 287-298.
- Maneerot, J., A. Terdwongworakul, W. Tanaphase, and N. Unprasert. 2006. Application of near infrared spectroscopy to predict crude protein in shrimp feed. *Kasetsart J. (Nat. Sci.)* 40: 172-180.
- Murray, I., L. S. Aucott, and I. H. Pike, 2001. Use of discriminant analysis on visible and near infrared reflectance spectra to detect adulteration of fish meal with meat and bone meal. *Near Infrared Spectrosc.* 9: 297-311.
- Pérez-Marín, D. C., A. Garrido-Varo, J. E. Guerrero-Ginel, and A. Gómez-Cabrera. 2004. Near-infrared reflectance spectroscopy (NIRS) for the mandatory labelling of compound feedingstuffs: chemical composition and open-declaration. *Anim. Feed Sci. Technol.* 116: 333-349.
- Reid, L. M., T. Woodcock, C. P. O'Donnell, J. D. Kelly, and G. Downey. 2005. Differentiation of apple juice samples on the basis of heat treatment and variety using chemometric analysis of MIR and NIR data. *Food Res. Int.* 38: 1109-1115.
- Roychoudhury, P., L. M. Harvey, and B. McNeil. 2006. At-line monitoring of ammonium, glucose, methyl oleate and biomass in a complex antibiotic fermentation process using attenuated total reflectance-mid-infrared (ATR-MIR) spectroscopy. *Anal. Chim. Acta* 561: 218-224.
- Shao, Y. N., Y. He, and C. Q. Wu. 2008. Dose detection of radiated rice by infrared spectroscopy and chemometrics. *J. Agric. Food Chem.* 56: 3960-3965.
- Suykens, J. A. K., and J. Vandewalle. 1999. Least squares support vector machine classifiers. *Neural Process. Lett.* 9: 293-300.
- Suykens, J. A. K., T. Van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle. 2002. *Least squares support vector machines*. Singapore: World Scientific.
- Sweeney, R. A.. 1989. Generic combustion method for determination of crude protein in feeds: collaborative study. *J. AOAC* 72: 770.
- van Raamsdonk, L. W. D., C. von Holst, V. Baeten, G. Berben, A. Boix, and J. de Jong. 2007. New developments in the detection and identification of processed animal proteins in feeds. *Anim. Feed Sci. Technol.* 133: 63-83.
- Wang, L., F. S. C. Lee, X. R. Wang, and Y. He. 2006. Feasibility study of quantifying and discriminating soybean oil adulteration in camellia oils by attenuated total reflectance MIR and fiber optic diffuse reflectance NIR. *Food Chem.* 95: 529-536.
- Wu, D., Y. He, J. H. Shi, and S. J. Feng. 2009. Exploring near and midinfrared spectroscopy to predict trace iron and zinc contents in powdered milk. *J. Agric. Food Chem.* 57: 1697-1704.
- Xiccato, G., A. Trocino, J. L. De Boever, L. Maertens, R. Carabaño, J. J. Pascual, J. M. Perez, T. Gidenne, and L. Falcao-E-Cunha. 2003. Prediction of chemical composition, nutritive value and ingredient composition of European compound feeds for rabbits by near infrared reflectance spectroscopy (NIRS). *Anim. Feed Sci. Technol.* 104: 153-168.
- Yan, Y. L., L. L. Zhao, D. H. Han, and S. M. Yang. 2005. *The foundation and application of near-infrared spectroscopy analysis*. Beijing, China: China Light Industry Press.

Yang, Z. L., L. J. Han, X. Liu, and Q. F. Li. 2008. Detecting and quantifying meat meal or meat and bone meal contamination in fishmeal by visible and near infrared reflectance spectra. *Anim. Feed Sci. Technol.* 147: 357-367.

Zhou, Q., S. Q. Sun, L. Yu, C. H. Xu, I. Noda, and X. R. Zhang. 2006. Sequential changes of main components in different kinds of milk powders using two-dimensional infrared correlation analysis. *J. Mol. Struct.* 799: 77-84.

## APPENDIX

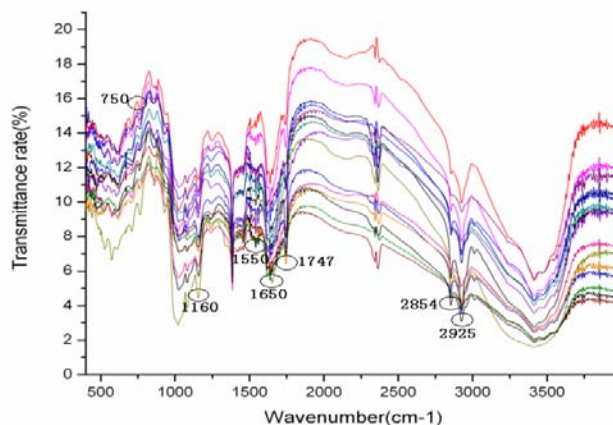


Figure 1. MIR transmission rate of fish feeds.

Table 1. Protein contents in calibration and prediction sets of fish feeds.

Set	Number of sample	Maximum (%)	Minimum (%)	Mean (%)	Standard deviation (%)
Total samples	225	37.2947	8.81543	25.2696	8.7534
Calibration set	150	37.2947	8.81543	25.2856	8.7603
Prediction set	75	36.8681	8.98719	25.2376	8.7395

Table 2. The performances of two models based on different spectral ranges.

Algorithm	Spectral range	Parameter		
		R	RMSEP (%)	Bias (%)
PLS	NIR	0.8147	5.0837	0.2057
	MIR	0.9907	1.1960	-0.0172
LS-SVM	NIR	0.8521	4.5782	0.3557
	MIR	0.9955	0.8918	0.3304