Application of NIR spectroscopy for tenderness classification of bovine meat

S. Lambotte[1], B. Leroy[2], O. Dotreppe[3], H. Lecocq[1], L. Istasse[3] and A. Clinquart[2].

[1] *University of Liège, Robotics and Automation Unit, Faculty of Applied Sciences, Belgium.*
[2] *University of Liège, Dept. of Food Science (Technology), Faculty of Veterinary Medicine, Belgium.*
[3] *University of Liège, Nutrition Unit, Faculty of Veterinary Medicine, Belgium.*

Key words: tenderness, near infrared spectroscopy, classification, Partial Least Squares, discrimant analysis, neural network.

Background

Tenderness is considered as the most important sensory attribute of meat, affecting consumer acceptance of beef. The determination of bovine meat tenderness with a rapid and non-destructive technique is the topic of several researchs (Belk *et al.*, 2000 ; Hildrum *et al.*, 1994). Among the fast and non-destructive techniques, the NIR spectroscopy seems to be promising. According to several previous studies, the Warner Bratzler Peak Shear Force (WBPSF) prediction by NIR spectroscopy did not seem to be accurate enough for routine application. For example, Byrne *et al.*(1998) obtained a standard error of prediction (SEP) of *2.1* kg for a standard deviation of 2.7 kg, Leroy *et al.*(2000) obtained a standard error of cross-validation (SECV) of *9.7* N for a standard deviation of *12.4* N.

Objectives

Alternatives to the prediction of WBPSF of bovine meat by usual regression techniques consist in classifications of beef cuts in several classes of tenderness. The objectives of this study were to explore and compare some mathematical methods to classify beef samples among classes (tender, intermediate and tough) from spectra acquired in reflection and transmission modes.

Material and Methods

Muscle *Longissimus thoracis* was removed from 190 bovine carcasses 2 days *post mortem*. A *2.5* cm thick heated cut was used to determine WBPSF value with a Lloyd LR5K universal testing machine perpendicular to the muscle fibre direction on ten *1.25* cm diameter cores. The NIR analysis was performed 2 days after slaughter with the Fourier Transform spectrometer Bomem MB 160D in the 800 to 2500 nm spectral range. The fiber optic Axiom probe was used for spectral acquisition in the reflection mode and the transmission mode was performed with the Bag SamplIR accessory. Five spectra were obtained in different places of the sample and the average spectrum was used to calculate the classification models. A principal component analysis (PCA) was first performed on the Grams/32 (Galactic) software to detect outliers. Preprocessings and models calculations were programmed in the Matlab (MathWorks) software, version 5.3. The preprocessings tested were the Baseline Correction (BLC), the Multiplicative Scattering Correction (MSC) and the 1° and 2° derivatives. The first of the five classification techniques was the Partial Least Squares regression (PLS) followed by discrimination. Linear Discriminant Analysis (LDA) is a classification method which calculates the distance between a sample and the gravity center of the different groups. Another classification technique calculates the distances from the other samples and determines the sample membership by the *k* nearest neighbours (KNN). The Probabilistic Neural Network (PNN) is a similar technique but the contribution of each sample is calculated. Artificial Neural Network (ANN), the last method used, is a regression technique based on an optimization which determines the parameters that minimize the prediction error of a nonlinear function. The PCA components were chosen for the LDA, KNN and PNN techniques. The PLS components were preferred for the techniques based on regression and discrimination (PLS and ANN). The discriminant values of WBPSF were arbitrary chosen equal to 42 and 62 N. A cross-validation was performed three times and the average of models performance (error rate) was examined to determine the parameters of the model and the number of components.

Results and discussion

Table 1 shows the best models for each classification technique on transmission and reflection modes. The transmission mode seemed to be better with an error rate from *33* to *43 %* *vs* *35* to *47 %* in reflection mode. It can be observed that the tender and tough samples were not well predicted. This could be explained by a smaller number of samples in these groups (*20* and *25 %* of the samples respectively in the tender and tough groups). To force a correct classification in the extreme groups (tender and tough), another model performance criterion than total error rate could be used. The Cumulative Error Rate (CER) is the sum of the error rate in each group. With this parameter as performance criterion, it is necessary to have a good prediction in each group and not only in the largest one. Table 2 gives the results using this performance criterion. Globally the prediction of the tender and the tough groups was improved but the prediction of the samples of intermediate tenderness was of lower quality. The models of classification among three groups did not give satisfactory results for routine check. Consequently, the classification among two groups was tested to detect only the toughest cuts. The discriminant value of WBPSF was chosen at *60* N. Table 3a shows the results of classification when the total error rate is used as performance criterion. In these conditions, the transmission mode (error rate of *19* to *22 %*) seemed more adapted than the reflection mode (error rate of *25* to *28 %*). The error in the group of toughest samples was higher (more than *42 %*) compared to the one of tender samples (less than *14 %*). Once again, the toughest samples represented a smaller group, *30 %* of the samples. The use of CER as performance criterion could improve the classification of the toughest samples. As showed in table 3b, classification of the toughest samples was better (in reflection mode, *48-56 %* using CER criterion *vs* *52-91 %* using error rate criterion) but remained of too weak quality. The error rate in the group of tough samples could be used as new performance criterion. Table 4 shows the best models for each classification method. The best detection of the toughest samples was realized with the PLS method in transmission mode. The error rate in the group of tough cuts was *37.5 %*, *15.1 %* in the tender samples group, the total error rate being *21.3 %*. These results did not allow a good classification of bovine tenderness. Nevertheless, these results can be compared with those of Belk *et al.* (2000) who used the *BeefCam* device to classify the samples in two groups of tenderness on the basis of color image analysis. They announced a total error rate of *38.8 %*, an error rate of *21.7 %* for the toughest samples and *41.5 %* for the tender samples. The *BeefCam* seemed to better detect the toughest samples but rejected too many tender samples. The difficulty to classify the samples could be caused by the heterogeneity of the meat. In fact, in the present study, the WBPSF value was obtained from *10* measurements on the same sample. To obtain a good prediction, it is important to work with a small standard error of the reference method (S_{REF}) compared with the standard deviation (SD) between samples. The S_{REF} obtained in the present study was equal to *9.1* N compared with a SD of *13.0* N. In these conditions, it is difficult to distinguish between tender and tough beef cuts.

Conclusions

Generally, the transmission acquisition mode was better than the reflection mode. The differences between classification models were not very large and the most complex techniques such as KNN, PNN or ANN did not bring a consistent improvement by comparison with PLS and LDA. To improve the quality of classification models, it would be imperative to reduce the intra-sample variability of the reference measurements.

	Classification méthod	Preprocessing	Number of components	Error rate of classification	Error rate between extrem groups	Error rate in the tender group	Error rate in the intermediate group	Error rate in the tough group
Reflection	*PLS*	*None*	*4*	*41.7*	*1.3*	*89.2*	*9.3*	*75.6*
	LDA	*BLC*	*14*	*46.8*	*2.8*	*85.3*	*20.0*	*67.9*
	KNN	*2° derivative*	*15*	*35.4*	*2.8*	*86.3*	*18.9*	*30.8*
	PNN	*1° derivative*	*17*	*45.0*	*1.9*	*83.3*	*26.3*	*61.5*
	ANN	*1° derivative*	*4*	*40.0*	*1.3*	*93.1*	*10.0*	*64.4*
Transmission	*PLS*	*MSC*	*4*	*34.1*	*0.2*	*65.7*	*14.3*	*56.7*
	LDA	*1° derivative*	*7*	*41.4*	*2.7*	*63.7*	*21.4*	*63.9*
	KNN	*None*	*5*	*30.8*	*1.7*	*85.3*	*14.9*	*59.0*
	PNN	*1° derivative*	*2*	*43.1*	*0.4*	*79.4*	*19.2*	*63.2*
	ANN	*1° derivative*	*9*	*35.1*	*1.5*	*67.6*	*15.0*	*49.2*

Table 1: Results of classification (cross-validation) of the samples among 3 groups (tender, intermediate, tough) – reflection and transmission modes – model performance criterion is the error rate of classification.

	Classification method	Preprocessing	Number of components	Cumulative error rate	Error rate of classification	Error rate between extrem groups	Error rate in the tender group	Error rate in the intermediate group	Error rate in the tough group
Reflection	*PLS*	*BLC*	*7*	*1.64*	*41.7*	*3.6*	*86.3*	*17.2*	*60.7*
	LDA	*2° derivative*	*2*	*1.53*	*57.4*	*10.4*	*28.4*	*70.0*	*54.5*
	KNN	*1° derivative*	*16*	*1.71*	*47.7*	*5.7*	*83.3*	*26.3*	*61.5*
	PNN	*1° derivative*	*15*	*1.63*	*51.9*	*9.7*	*59.8*	*45.6*	*57.7*
	ANN	*1° derivative*	*20*	*1.36*	*49.1*	*4.5*	*66.7*	*45.4*	*43.7*
Transmission	*PLS*	*1° derivative*	*6*	*1.26*	*35.4*	*0.8*	*59.8*	*18.0*	*48.3*
	LDA	*None*	*2*	*1.35*	*48.5*	*5.6*	*40.2*	*57.6*	*36.8*
	KNN	*None*	*3*	*1.54*	*42.0*	*3.4*	*78.4*	*22.8*	*52.8*
	PNN	*2° derivative*	*20*	*1.49*	*49.6*	*6.9*	*52.0*	*50.4*	*46.5*
	ANN	*None*	*10*	*1.25*	*33.5*	*1.0*	*58.8*	*19.7*	*46.7*

Table 2: Results of classification (cross-validation) of the samples among 3 groups (tender, intermediate, tough) – reflection and transmission modes – model performance criterion is the cumulative error rate.

	Classification method	Preprocessing	Number of components	Error rate of classification	Error rate in the tender group	Error rate in the tough group
Reflection	*PLS*	*1° derivative*	*6*	*25.4*	*14.2*	*51.9*
	LDA	*None*	*15*	*25.8*	*4.6*	*76.3*
	KNN	*2° derivative*	*16*	*27.7*	*5.6*	*80.1*
	PNN	*2° derivative*	*17*	*27.7*	*1.1*	*91.0*
	ANN	*None*	*7*	*28.3*	*11.6*	*61.5*
Transmission	*PLS*	*None*	*3*	*19.5*	*9.0*	*46.5*
	LDA	*MSC*	*9*	*22.4*	*6.3*	*64.6*
	KNN	*MSC*	*6*	*21.3*	*7.1*	*59.7*
	PNN	*MSC*	*14*	*22.0*	*7.1*	*61.1*
	ANN	*1° derivative*	*9*	*19.9*	*11.4*	*42.4*

Table 3a: Results of classification (cross-validation) of the samples among 2 groups (tender, tough) – reflection and transmission modes – model performance criterion is the error rate of classification.

Preprocessing	Number of components	Cumulative error rate	Error rate of classification	Error rate in the tender group	Error rate in the tough group
1° derivative	*6*	*0.66*	*25.4*	*14.2*	*51.9*
None	*6*	*0.71*	*27.5*	*15.6*	*55.8*
1° derivative	*13*	*0.71*	*28.2*	*17.7*	*53.2*
None	*2*	*0.72*	*30.9*	*23.7*	*48.1*
BLC	*12*	*0.69*	*27.8*	*18.5*	*50.0*
None	*9*	*0.55*	*21.3*	*15.1*	*37.5*
None	*2*	*0.59*	*26.4*	*22.5*	*36.8*
MSC	*3*	*0.63*	*22.0*	*10.3*	*52.8*
2° derivative	*15*	*0.65*	*27.4*	*22.5*	*40.3*
1° derivative	*9*	*0.54*	*19.9*	*11.4*	*42.4*

Table 3b: Results of classification (cross-validation) of the samples among 2 groups (tender, tough) – reflection and transmission modes – model performance criterion is the cumulative error rate.

	Classification method	Preprocessing	Number of components	Error rate of classification	Error rate in the tender group	Error rate in the tough group
Reflection	*PLS*	*None*	*20*	*30.7*	*24.5*	*45.5*
	LDA	*2° derivative*	*2*	*31.1*	*22.6*	*51.5*
	KNN	*None*	*2*	*34.1*	*26.1*	*53.2*
	PNN	*None*	*2*	*30.9*	*23.7*	*48.1*
	ANN	*1° derivative*	*18*	*32.8*	*27.2*	*46.5*
Transmission	*PLS*	*None*	*9*	*21.3*	*15.1*	*37.5*
	LDA	*None*	*2*	*26.4*	*22.5*	*36.8*
	KNN	*2° derivative*	*14*	*32.0*	*29.1*	*39.8*
	PNN	*2° derivative*	*15*	*27.4*	*22.5*	*40.3*
	ANN	*1° derivative*	*20*	*29.9*	*25.7*	*41.0*

Table 4: Results of classification (cross-validation) of the samples among 2 groups (tender, tough) – reflection and transmission modes – model performance criterion is the error rate of tough samples.

References

Belk, K., Scanga, J.A., Wyle, A.M., Wulf, D.M., Talum, J.D., Reagan J.O., & Smith G.C. (2000). The use of Video Image Analysis and Instrumentation to predict beef palatability. *53rd Annual Reciprocal Meat Conference*, 10-15.

Byrne, C.E., Downey, G., Troy, D.J., & Buckley, D.J. (1998). Non-destructive prediction of selected quality attributes of beef by near infrared reflectance spectroscopy between 750 and 1098 nm. *Meat Science*, 49 (4), 399-409.

Hildrum, K.I., Nilsen, B.N., Mielnik, M., & Naes, T.(1994). Prediction of sensory characteristics of beef by near infrared spectroscopy. *Meat Science*, 38, 67-80.

Leroy, B., Lambotte, S., Dotreppe, O., Lecocq, H., Istasse, L. ., & Clinquart, A., 2000. Prediction of technological and organoleptic properties of beef cuts by Near Infrared Spectroscopy. *Proceedings of the 46th International Congress of Meat Science and Technology*, Buenos Aires, 610-611.

This work is part of a project supported by the Direction Générale des Technologies, de la Recherche et de l'Energie, Région Wallonne, Belgium. Project N°981/3723.