# Near infrared spectroscopy for cow milk composition analysis: the influence of mastitis

**Roumiana Tsenkova,[a] Stefka Atanassova,[b] Yukihiro Ozaki,[c] Kazuhiko Itoh[d] and Kiohiko Toyoda[a]**

[a]*Kobe University, Agricultural Faculty, Kobe 657, Japan.*

[b]*Thracian University, Agricultural Faculty, 6000 Stara Zagora, Bulgaria.*

[c]*Kwansei-Gakuin University, School of Science, Nishinomia 662, Japan.*

[d]*Hokkaido University, Faculty of Agriculture, Sapporo 060, Japan.*

## Introduction

Measurement of milk composition is essential for the dairy industry and the management of dairy farms. The quality of milk has a direct influence on the quality of processed milk products. Daily measurement of every individual cow's milk composition is important for animal breeding, efficient utilisation of cows and for nutrition management.[1,2] Mastitis is an intramammary bacterial infection, which is common disease among dairy cattle.[3,4] It causes a decrease in milk yield. Chemical changes in milk composition, due to mastitis, reduce milk quality.

Near infrared (NIR) spectroscopy has been applied to fat, protein and lactose measurement of raw milk.[5–11] The purpose of this study was to investigate the influence of mastitic milk samples in the data set used for the development of milk composition calibration equations in the NIR region from 1100 to 2400 nm on the accuracy of NIR determination of fat, protein and lactose content of non-homogenised milk.

## Material and methods

A total of 260 individual milk samples were collected from three Holstein cows. The samplings were done twice a month, for a six month period, every second and fourth week, at morning and evening milking, respectively. Teat cups were removed at each three litres of milk yield as determined by fractional sampling milk meter True-Test (Alfa Laval Agri, Tumba, Sweden). The experiment was carried out in the period between the second and eighth months of lactation of the examined cows. All cows were three years old, in their second lactation and on the same diet.

All samples were analysed for fat, total protein and lactose by Milko-scan (Foss-Electric A/C, Hillerød, Denmark)[12] and for somatic cell count (SCC) by Foss-somatic (Foss-Electric A/C, Hillerød, Denmark). Thirty-five milk samples had a somatic cell count (SCC) higher than 200 000SCC cell $ml^{-1}$, which has been accepted as an indicator for mastitis.

Transmittance ($T$) milk spectra with a 1 mm sample thickness were obtained from all samples by NIRSystem 6500 spectrophotometer (Foss NIRSysterns, Silver Spring, MD, USA). Spectra were recorded in the wavelength range from 1100 to 2400 nm at 2 nm intervals and were stored in the linked

computer as absorbance, i.e. log (1/*T*). Prior to spectral analysis each sample was warmed up to 40°C in a water bath.

A commercial program, Pirouette Version 2.0 (Infometrics, Inc., Woodinville, WA, USA), was used to process the data and to develop models for fat, total protein and lactose content determination. The set of all samples and the set of samples with low SCC were randomly divided into a calibration sub-set (two-thirds of the samples) and test sub-set (one-third of the samples), respectively. Calibration equations were obtained for different sample sets: calibration set from all samples; calibration set of samples with low SCC (SCC < 200 000 cell $ml^{-1}$); and set of samples with high SCC (SCC > 200 000 cell $ml^{-1}$).

The methods for data treatment included first derivative transformation of log (1/*T*) data with window size 25 point, based on Savitzki–Golay polynomial filter.

The calibration was performed using partial least square (PLS) regression. In the development of all calibration models, ten PLS factors were set up as a maximum. The optimum number of PLS factors used in the models was determined by a cross-validation method. In cross-validation, five samples were temporarily removed from the calibration set to be used for validation. With the rest of the samples a PLS model was developed and applied to predict the fat content of the group of five samples. The results were compared to the respective reference values. This procedure was repeated several times until a prediction for all samples was obtained. Performance statistics were accumulated for each group of removed samples. The validation errors were combined into a standard error of cross-validation (*SECV*). The optimum number of PLS factors in each model was defined to be the one that corresponded to the lowest *SECV.*

Calibration and validation statistics for each calibration equation, standard error of calibration (*SEC*), coefficient of multiple correlation (*R*), standard error of prediction (*SEP*), correlation coefficient between measured and NIR predicted value (*r*), bias and skew were calculated and were used for evaluation of the NIR determination accuracy.

**Table 1. Mean, range and standard deviation (SD) for milk samples of the different calibration and test sets used.**

| Sample set | Parameter | Average | Min. | Max. | SD |
|---|---|---|---|---|---|
| Low SCC: calibration set | Fat, % | 3.18 | 0.47 | 8.35 | 1.72 |
| (*n* = 149) | Protein, % | 3.08 | 2.60 | 3.62 | 0.19 |
| | Lactose, % | 4.41 | 4.08 | 4.74 | 0.15 |
| Low SCC: test set | Fat, % | 3.24 | 0.69 | 6.71 | 1.62 |
| (*n* = 74) | Protein, % | 3.08 | 2.60 | 3.61 | 1.92 |
| | Lactose, % | 4.40 | 4.09 | 4.71 | 0.14 |
| High SCC set | Fat, % | 4.89 | 0.34 | 10.41 | 2.41 |
| (*n* = 35) | Protein, % | 3.10 | 2.84 | 3.39 | 0.15 |
| | Lactose, % | 4.39 | 4.12 | 4.70 | 0.16 |
| All samples: calibration set | Fat, % | 3.51 | 0.34 | 10.41 | 1.90 |
| (*n* = 172) | Protein, % | 3.13 | 2.70 | 3.99 | 0.24 |
| | Lactose, % | 4.40 | 4.09 | 4.74 | 0.17 |
| All samples: test set | Fat, % | 3.50 | 0.65 | 9.10 | 1.98 |
| (*n* = 86) | Protein, % | 3.13 | 2.60 | 3.62 | 0.22 |
| | Lactose, % | 4.40 | 4.08 | 4.74 | 0.16 |

**Table 2. NIR calibration and validation statistics for fat content determination in different milk sample sets.**

| Calibration set | PLS factors | *SEC* | *R* | Test set | *SEP* | *r* | Bias | Skew |
|---|---|---|---|---|---|---|---|---|
| low SCC | 9 | 0.091 | 0.999 | low SCC | 0.118 | 0.997 | 0.0003 | 0.999 |
| low SCC | 9 | 0.091 | 0.999 | high SCC | 0.687 | 0.968 | 0.047 | 0.985 |
| high SCC | 8 | 0.042 | 0.999 | low SCC | 0.158 | 0.996 | 0.021 | 0.988 |
| all samples | 10 | 0.080 | 0.999 | all samples | 0.123 | 0.998 | 0.012 | 0.998 |

**Table 3. NIR calibration and validation statistics for total protein content determination in different milk sample sets.**

| Calibration set | PLS factors | *SEC* | *R* | Test set | *SEP* | *r* | Bias | Skew |
|---|---|---|---|---|---|---|---|---|
| low SCC | 10 | 0.090 | 0.882 | low SCC | 0.100 | 0.853 | 0.0007 | 1.012 |
| low SCC | 10 | 0.090 | 0.882 | high SCC | 0.820 | 0.207 | 0.096 | 0.050 |
| high SCC | 4 | 0.092 | 0.829 | low SCC | 0.198 | 0.192 | -0.060 | 0.479 |
| all samples | 10 | 0.097 | 0.855 | all samples | 0.119 | 0.850 | 0.010 | 0.914 |

**Table 4. NIR calibration and validation statistics for lactose content determination in different milk sample sets**

| Calibration set | PLS factors | *SEC* | *R* | Test set | *SEP* | *r* | Bias | Skew |
|---|---|---|---|---|---|---|---|---|
| low SCC | 10 | 0.074 | 0.871 | low SCC | 0.079 | 0.834 | -0.0047 | 1.011 |
| low SCC | 10 | 0.074 | 0.871 | high SCC | 1.417 | 0.271 | 0.027 | 0.747 |
| high SCC | 2 | 0.115 | 0.696 | low SCC | 0.145 | 0.425 | -0.060 | 0.788 |
| all samples | 10 | 0.086 | 0.872 | all samples | 0.079 | 0.847 | 0.001 | 0.861 |

## Results and discussion

The ranges, mean values and standard deviation of fat, total protein and lactose content of milk samples from different sample sets, resulting from standard chemical analysis, were presented in Table 1. The calibration equations for fat, total protein and lactose content determination for the different calibration sets and verification statistics obtained using different test sets of samples are shown in Tables 2 to 4.

The NIR calibrations for fat, total protein and lactose content for the set of samples with low SCC were very successful. The obtained *SEP* was close to the accuracy of the respective reference methods.[12] The bias was very small and skew was in the range from 0.999 to 1,012. The accuracy of calibration models for the set of samples with low SCC, when compared to the set of all samples, showed a lower *SEP* for fat and total protein content determination and an equal one for lactose.

The calibration equations, developed from the set with high SCC milk samples had a *SEC* lower for fat content determination, similar for total protein content and larger for lactose content, respectively, when compared to the *SEC* for other used sample sets.

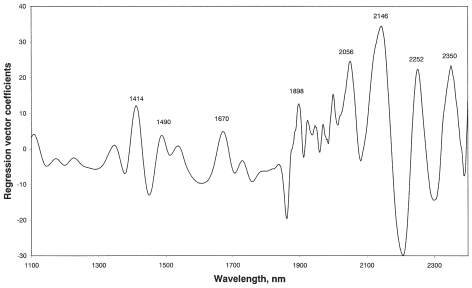


**Figure 1. Regression vector for protein determination in the set of samples with low somatic cell count.**

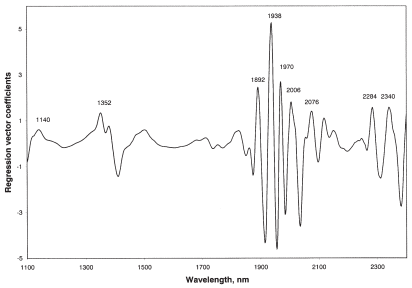


**Figure 2. Regression vector for protein determination in the set of samples with high somatic cell count.**

The *SEP* and bias increased, correlation coefficient decreased and skew changed dramatically when equations for low SCC milk were used to predict fat, total protein and lactose content in the data set including samples with high SCC. The calibration equations, derived using samples with low SCC content, were not suitable for determination of examined components in the data set contained milk with high level of SCC, especially for protein and lactose determination. A similar result was given when the models obtained for the set with high SCC milk samples were used to measure fat, total protein and lactose content in the set of samples with low SCC. The presence of milk samples with high SCC in the data set used for calibration or prediction highly influenced the accuracy of fat, protein and lactose determination.

The mean value and the range of fat, total protein and lactose content in the calibration and test set of samples with low SCC and the set of samples with high SCC were approximately the same. Therefore, the impossibility of applying calibration equations, developed from the set with low SCC milk for determination of fat, total protein and lactose content in the set containing milk with high SCC was not connected to differences in the range of measured constituents. A possible reason for this effect is the chemical changes in milk, caused by mastitis. Mastitis decreases lactose content.[2,3] Although total protein content of milk undergoes little change, the types of proteins change dramatically. It is known that mastitis causes a rise in whey protein and a decrease in casein. Mastitis also changes the ionic concentration in the milk. Alterations to the blood–milk barrier lead to an increase in sodium chloride and a decrease in potassium in the milk. These changes in milk composition influenced NIR spectral data and were confirmed by the differences in regression vectors. For example, there were significant differences in the regression vectors for total protein determination for the set of low SCC and the set of high SCC milk (Figures 1 and 2). The biggest regression coefficients in the equations for protein determination in the high SCC sample set were found for the spectral region from 1890 to 2033 nm, which was not observed in the models for protein determination in the set of milk with low SCC. The main absorbance bands of milk in this region were connected with water and protein absorption.[13] A possible reason for that significant influence could be an increasing quantity of water-soluble proteins in the milk and changes in ionic concentration.[2,3]

## Conclusion

It has been found that mastitic milk samples with high SCC in data set used for calibration or prediction highly influenced the accuracy of NIR fat, protein and lactose content determination. It is necessary to take this influence into consideration when calibration models for the determination of milk chemical composition are developed. There are two possible approaches to avoid incorrect determina-

tion of milk composition. One is to develop an initial classification of milk samples according to their SCC and using respective calibration equations for each group. Another one, with little compromise concerning accuracy, is to use the calibration models, obtained from the calibration set of milk samples with different levels of SCC.

## Acknowledgments

This work was supported by the Programme for Promotion of Basic Research Activities for Innovative Bioscience (PROBRAIN), Ministry of Agriculture, Fisheries and Food, Japan. The authors are grateful to GL Science Inc., Tokyo, Japan for providing the Pirouette Version 2.0 software.

## References

1. K. Svennersten-Sjaunja, L.O. Sjaunja, J. Bertilsson and H. Wiktorsson, *Livestock Prod. Sci.* **48,** 167 (1997).
2. J. Hamann and V. Krornker, *Livestock Prod. Sci.* **48,** 201 (1997).
3. O.W. Shalm, E.J. Carrol and N.C. Jain, *Bovine Mastitis*. Lea and Febiger, Philadelphia, USA (1971).
4. R.J. Harmon, *J. Dairy Sci.* **77,** 2103 (1994).
5. T. Sato, M. Yoshino, S. Farukawa, Y. Someya, N. Yano, J. Uozumi and M. Iwamoto, *Jpn J. Zootech. Sci.* **58,** 698 (1987).
6. R. Tsenkova and T. Grigorov, *Farm Machinery* **27,** 81 (1990).
7. Z. Schmilovitch, E. Maltz and M. Austerweil, in *Prospects for Automatic Milking*. EAAP Publ. No. 65. PUDOC Scientific Publishing, Wageningen, The Netherlands, p. 193 (1992).
8. R.N. Tsenkova, K.I.Yordanov and Y. Shinde, in *Prospects for Automatic Milking*. EAAP Publ. No. 65. PUDOC Scientific Publishing, Wageningen, The Netherlands, p. 185 (1992).
9. E. Diaz-Carrillo, A. Munoz-Serrano, A. Alonso-Moraga and J.M. Serradilla-Manrique, *J. Near Infrared Spectrosc*. **1,** 141 (1993).
10. R.N. Tsenkova, K.I. Yordanov, K. Itoh, Y. Shindle and J. Nishibu, in *Dairy System for the 21st Century.* Proc. 3rd Int. Dairy Housing Conference. ASAE, St. Joseph, Michigan, USA, p. 82 (1994).
11. R.N. Tsenkova, K. Itoh, J. Himoto and K. Asahida, in *Leaping Ahead with Near Infrared Spectroscopy*, Ed by G.D. Batten, P.C. Flinn, L.A. Welsh and A.B. Blakeney. NIR Spectrosc. Group, Melbourne, Australia, p. 329 (1995).
12. Official Methods ofAnalysis, 15th ed. AOAC, Arlington, VA, USA (1990).
13. B. Osborne, T. Fearn and P.H. Hindle, *Practical NIR Spectroscopy with Application in Food and Beverage Analysis*. Longman Scientific and Technical, Harlow, UK (1993).