

## Original article

# Mid-infrared spectrometry: A tool for the determination of chemical parameters in Emmental cheeses produced during winter

Romdhane KAROUI<sup>a\*</sup>, Abdul Mounem MOUAZEN<sup>a</sup>, Éric DUFOUR<sup>b</sup>, Laurent PILLONEL<sup>c</sup>, Daniel PICQUE<sup>d</sup>, Jacques-Olivier BOSSET<sup>c</sup>, Josse DE BAERDEMAEKER<sup>a</sup>

<sup>a</sup> Division of Mechatronics, Biostatistics and Sensors (MeBioS), Department of Biosystems, K.U. Leuven, Kasteelpark Arenberg 30, 3001, Leuven, Belgium

<sup>b</sup> U.R. “Typicité des Produits Alimentaires”, ENITA de Clermont-Ferrand, BP 35, 63370 Lempdes, France

<sup>c</sup> Agroscope Liebefeld-Posieux (ALP), 3003 Berne, Switzerland

<sup>d</sup> Laboratoire de Génie et Microbiologie des Procédés Alimentaires, Institut National de la Recherche Agronomique, 78850 Thiverval Grignon, France

Received 18 May 2005 – Accepted 11 October 2005

**Abstract** – The study investigates the potential use of the mid-infrared (MIR) spectroscopy in transmission mode for determining some chemical parameters of European Emmental cheeses. A total of 91 Emmental cheeses produced during the winter in Austria ( $n = 4$ ), Finland ( $n = 6$ ), Germany ( $n = 13$ ), France ( $n = 30$ ) and Switzerland ( $n = 38$ ) were investigated. Fat, total nitrogen (TN), water-soluble nitrogen (WSN), non-protein nitrogen (NPN), sodium chloride (NaCl) and pH were analysed by the reference methods. The MIR transmission of different cheeses was also measured by a Nicolet Magna 750 IR spectrometer in a measurement range of 4000–600 cm<sup>-1</sup>. Partial least square analysis together with the leave-one-out cross-validation technique was used to develop calibration models for different cheese parameters. The best results for NPN ( $R^2 = 0.85$ ; ratio of standard deviation to root mean square error of prediction (RPD) = 2.58), pH ( $R^2 = 0.82$ ; RPD = 2.34), WSN ( $R^2 = 0.80$ ; RPD = 2.26), NaCl ( $R^2 = 0.70$ ; RPD = 1.91), fat ( $R^2 = 0.69$ ; RPD = 1.87), TN ( $R^2 = 0.62$ ; RPD = 1.60), WSN/TN ( $R^2 = 0.83$ ; RPD = 2.81) and WSN/NPN ( $R^2 = 0.57$ ; RPD = 1.52) were obtained when the spectra were subjected to the first derivation and smoothing after maximum normalisation. It can be concluded that MIR-transmission spectroscopy can be considered as a suitable technique for the prediction of NPN, pH, WSN and WSN/TN of European Emmental cheeses with different origins. The NaCl, TN, WSN/NPN and fat can also be estimated, but with much lower precision.

**Emmental cheese / mid-infrared spectroscopy / calibration / partial least square method**

**摘要 – 中红外光谱法测定冬季生产 Emmental 干酪的化学参数。**本文尝试采用中红外光谱法 (MIR) 测定欧洲生产的 Emmental 干酪中一些化学参数变化。本研究共选择了不同国家和不同乳品厂在冬季生产的 91 份 Emmental (瑞士多孔) 干酪样品，其中奥地利 4 份、芬兰

\* Corresponding author (通讯作者): Romdhane.Karoui@biw.kuleuven.be

6份、德国13份、法国30份、瑞士38份。脂肪、总氮(TN)、水溶性的氮(WSN)、非蛋白质氮(NPN)、氯化钠(NaCl)和pH值分别按参考文献给定的方法分析，同时采用Nicolet Magna 750红外光谱仪测定干酪中上述组分，扫描波数范围为4000–600 cm<sup>-1</sup>。采用偏最小二乘法(PLS)和留一法交互验证技术对不同的干酪参数建立了校正模型。

将数据经规范化处理后的最佳结果为NPN ( $R^2 = 0.85$ , RPD = 2.58, RPD为标准偏差与预测集均方根误差的比值)、pH ( $R^2 = 0.82$ , RPD = 2.34)、WSN ( $R^2 = 0.80$ , RPD = 2.26)、NaC ( $R^2 = 0.70$ , RPD = 1.91)、脂肪 ( $R^2 = 0.69$ , RPD = 1.87)、TN ( $R^2 = 0.62$ , RPD = 1.60)、WSN/TN ( $R^2 = 0.83$ , RPD = 2.81) 和 WSN/NPN ( $R^2 = 0.57$ , RPD = 1.52)。因此，可以得出结论，中红外光谱法适合于欧洲不同地区生产的Emmental干酪中NPN、pH、WSN和WSN/TN含量的预测，也可估算出干酪样品中NaCl、TN、WSN/NPN和脂肪的含量，但精密度较差。

### **Emmental 干酪 / 中红外光谱 / 校正 / 偏最小二乘法**

**Résumé – La spectroscopie moyen infrarouge : un outil pour la détermination des paramètres chimiques des fromages Emmental de production hivernale.** Cette étude traite du potentiel d'utilisation de la spectroscopie moyen infrarouge en mode transmission pour la détermination de certains paramètres chimiques des fromages Emmental européens. Un total de 91 fromages Emmental de production hivernale, provenant de : Autriche ( $n = 4$ ), Finlande ( $n = 6$ ), Allemagne ( $n = 13$ ), France ( $n = 30$ ) et Suisse ( $n = 38$ ) ont été analysés. La matière grasse, l'azote total, l'azote soluble, l'azote non protéique, le chlorure de sodium et le pH ont été déterminés par les méthodes de références. La spectroscopie moyen infrarouge en mode transmission a été, également, appliquée aux différents fromages en utilisant un spectromètre (Nicolet Magna 750) et ce, entre 4000–600 cm<sup>-1</sup>. La méthode aux moindres carrés partiels avec validation croisée a été utilisée pour développer des modèles de calibration, et ce pour tous les paramètres chimiques. Les meilleurs résultats pour l'azote non protéique ( $R^2 = 0,85$ ; ratio de la déviation moyenne à la racine carré de prédiction (RPD) = 2,58), pH ( $R^2 = 0,82$ ; RPD = 2,34), azote soluble ( $R^2 = 0,80$ ; RPD = 2,26), chlorure de sodium ( $R^2 = 0,70$ ; RPD = 1,91), matière grasse ( $R^2 = 0,69$ ; RPD = 1,87), azote total ( $R^2 = 0,62$ ; RPD = 1,60), rapport azote soluble/azote total ( $R^2 = 0,83$ ; RPD = 2,81) et du rapport azote soluble/azote non protéique ( $R^2 = 0,57$ ; RPD = 1,52) ont été obtenus lorsque la première dérivée a été appliquée aux spectres normés et lissés. La spectroscopie moyen infrarouge en mode transmission peut être considérée comme une technique appropriée pour la détermination des teneurs en azote non protéique, azote soluble, du rapport azote soluble/azote total et du pH des fromages Emmental européens. Les teneurs en NaCl, azote total, du rapport azote soluble/azote non protéique et en matière grasse peuvent être également prédictes, mais avec une précision beaucoup faible.

### **Emmental / spectroscopie moyen infrarouge / modèle de calibration / méthode des moindres carrés partiels**

## **1. INTRODUCTION**

In common with the processed food industry at large, the dairy industry has come under increasing pressure to deliver products of high and constant quality to the marketplace. Emmental cheese is a widespread cheese type manufactured in almost all industrial countries. Its added value depends on the technology used [34]. The production of Emmental cheeses with high quality requires monitoring of the product through a good cheese-making procedure

as well as good storage and maturation facilities. During the ripening phase, cheese undergoes a complex series of chemical, bacterial and enzymatic reactions. Concerning the biochemical changes that occur during ripening, proteolysis is the most important one owing to its major effect on texture and flavour development. During the ripening time, proteins are degraded into so-called primary products such as polypeptides and subsequently into secondary products such as small- and medium-size peptides, and eventually into free amino

acids [8]. The determination of the main constituents of cheese is important in the dairy food industry, to establish the cheese value, for consumers' information and for quality control. Therefore, routine series of physico-chemical analysis are required, although they are costly, time-consuming and require experienced operators.

In the last few decades, there has been a quest to replace traditional time-consuming reference methods such as Kjeldahl for protein, potentiometric titration for sodium chloride, etc. by spectroscopic techniques. Among these spectroscopic techniques, infrared (IR) represents a very promising tool. Near-infrared (NIR) spectrometry as a technique used for on-line and off-line analyses in food industries has been reviewed [35]. In recent years, NIR has become popular for process monitoring in many industries such as the dairy industry, in part because of the relative simplicity of the equipment required to record a spectrum [1, 2, 4, 7, 15, 23, 25, 29, 38, 41]. In addition, the relatively low absorption coefficients exhibited by many species in the NIR region implies that bulk samples can often be examined by NIR directly without the need for dilution in some inert, non-absorbing medium, as might be required for study in the fundamental mid-infrared (MIR) region.

The MIR wavelength range of 4000 to 400 cm<sup>-1</sup> has particular attraction, since measurements in this range provide direct information concerning the specific constituents in the sample, as well as their characteristic molecular structure [11]. MIR has been used for the determination of protein concentration in milk [11] and some chemical cheese parameters such as pH, total nitrogen (TN), fat, water-soluble nitrogen (WSN) and non-protein nitrogen (NPN) [5, 11, 28]. In addition, it has been reported that the MIR spectrum recorded on a cheese is a fingerprint of the sample that contains information about its physico-chemical characteristics [13, 14, 24]. Recently, both MIR and front-face fluorescence spectroscopies have been used for the characterisation of cheese structure [9, 26, 27] and for the determination of the geographic origin of PDO Gruyère and L'Etivaz PDO cheeses [10, 21] and of Emmental cheeses from dif-

ferent European geographic origins [18–20, 22, 33].

MIR spectroscopy is a method used for milk and dairy product analysis. Indeed, each chemical substance, apart from some salts and very simple chemical compounds, has its own distinctive spectrum. Only substances with very low concentrations (below 0.1%) can be difficult to determine, as noise might be encountered in such a case. Most of the absorption bands in the MIR region, but not in the NIR region, have been identified and attributed to chemical groups [21]. MIR spectra collected from a complex food product such as cheese cannot be interpreted using simple linear regression models, since the spectra may contain broad overlapping bands of various components. Therefore, to extract relevant information from the spectral data, a multivariate model is more appropriate and useful. Typically, multivariate statistical analysis such as partial least square (PLS) regression is used to relate variation in the spectra to the chemical parameters of cheeses. Thus, the objective of the current study was to evaluate the feasibility of using MIR in a transmittance mode to predict some chemical parameters in European Emmental cheeses with different origins.

## 2. MATERIALS AND METHODS

### 2.1. Origin and selection of the cheese samples

The geographic origin and the age of the Emmental cheeses produced during winter time have been reported previously by Karoui et al. [22]. The ripening time (at point of sampling) of the selected samples varied greatly, depending on their geographic origin. According to manufacturing usage, Emmental may be sold up to a maximum of 6–7 weeks of ripening or after a minimum of 3 months (Switzerland). The samples were selected based on their ripening time to reflect what consumers may expect to find in the store. The scattering of some results within a given region was partly explained by the diversity in the ripening time of the Emmental samples available (from young to very mature). In Switzerland,

for instance, it was very difficult to find Emmental ripened for only 90 days. Most of the Swiss samples had a ripening time varying from 4 to 6 months, with some of them reaching 12 months or even more.

The numbers and origins of the Emmental cheeses considered in this study were 4 from Austria, 6 from Finland, 13 from Germany, 30 from France and 38 from Switzerland. The cheeses from Austria, Germany and Switzerland were made from raw milk, whereas the cheeses from Finland were made from thermised milk. The cheeses from France were made up of two groups: the first group ( $n_1=15$ ) was made from raw milk and the second group ( $n_2=15$ ) was made from thermised milk (63 °C, 30 s). For each country, the samples were produced by different dairy plants, except for the Finnish, Austrian and French cheeses. In fact, 3 Finnish cheeses were manufactured by one dairy plant and the other 3 by another dairy plant. Only 2 Austrian and 2 French cheeses were made in the same dairy plant in Austria and France, respectively.

## 2.2. Physico-chemical analysis

The fat content was determined according to Gerber van Gulik [17], whereas total nitrogen (TN), water-soluble nitrogen (WSN) and non-protein nitrogen (NPN) were determined by the Kjeldahl method [6]. Sodium chloride (NaCl) was determined with potentiometric titration using a silver electrode [16]. The pH value was determined at room temperature using a penetrometric glass electrode (Mettler-Toledo, No. 104063123; AG, Im Langacher, 8606 Greifensee, Switzerland).

## 2.3. Mid-infrared spectroscopy

Five grams of fresh samples were dispersed in 100 mL water with a Polytron PT2100 at 15000 rpm for 60 s. An aliquot of 60 µL was applied to a polyethylene card (Spectra Tech Nicolet) used in the spectral range of 4000 cm<sup>-1</sup> to 600 cm<sup>-1</sup>, except between 2918 cm<sup>-1</sup> and 2849 cm<sup>-1</sup>, the region of absorption of the polyethylene card. The samples were dried for one night.

The prepared cards were analysed in transmittance mode using a Nicolet Magna 750 IR spectrometer in the measurement range between 4000 and 600 cm<sup>-1</sup>. Thirty-two interferograms with a resolution of 4 cm<sup>-1</sup> were recorded and averaged. For each cheese, three different aliquots prepared and three different cheese samples were analysed.

## 2.4. Mathematical treatment of data

### 2.4.1. Processing of spectra

The number of variables (wavelength) was first reduced to one-fourth. The cheese spectra were then normalised by the maximum normalisation technique. Normalisation is typically used to get all data to approximately the same scale, or to get a more even distribution of the variances and the average values. The maximum normalisation process is performed on a sample spectrum by dividing each transmission value by the sample maximum absolute transmission, as follows:

$$X'(i, k) = X(i, k) / \text{Max}(\text{Abs}(X(i, *))) \quad (1)$$

where  $X'(i, k)$  is the maximum normalised spectrum,  $X(i, k)$  is the sample transmission (%) and  $X(i, *)$  is the sample maximum absolute transmission (%).

The maximum normalisation is a normalisation that “polarises” the spectra. The peaks of all spectra with positive values touch +1, while spectra with values of both signs touch -1. Since all the cheese spectra in this study had positive values, the peaks of these spectra touched +1.

The first derivative of the normalised spectra was computed using the Savitzky-Golay method. This method enables the computation of the first- or higher-order derivatives, including a smoothing factor, which determines how many adjacent variables will be used to estimate the polynomial approximation used for derivation. A second-order polynomial approximation was selected. The first derivative was used for fat, NaCl, NPN, pH, TN, WSN, WSN/TN and WSN/NPN followed by 2:2

**Table I.** Sample statistics of calibration set used for PLS-cross-validation models<sup>a</sup>.

Compositional parameter	Minimum	Maximum	Mean	Standard deviation
Fat (g·kg <sup>-1</sup> )	280.00	360.00	313.56	14.34
NaCl (g·kg <sup>-1</sup> )	1.90	13.60	4.73	2.29
pH	5.49	5.95	5.66	0.09
NPN (g·kg <sup>-1</sup> )	2.02	9.82	6.51	1.40
TN (g·kg <sup>-1</sup> )	40.14	47.03	44.46	1.34
WSN (g·kg <sup>-1</sup> )	3.74	13.27	9.38	1.95
WSN/TN	0.08	0.30	0.21	0.05
WSN/NPN	1.02	2.36	1.44	0.18

<sup>a</sup> Abbreviation: NPN, non-protein nitrogen; TN, total nitrogen; WSN, water-soluble nitrogen.

smoothing to eliminate noise from the spectra, which was the final stage before model establishment.

#### 2.4.2. Model establishment

The objective of this task is to build a statistical model between the processed spectra and cheese properties. Several statistical modelling techniques can be adopted for proper calibration performance, such as linear and non-linear multiple regression analyses, principal component regression (PCR) and the PLS. The PLS was used in this study to relate the variations in one response variables (fat, pH, NaCl, NPN, TN, WSN, WSN/TN and WSN/NPN) to the variations of several predictors (wavelengths). The 273 spectral measurements were divided into two groups of 244 (calibration) and 29 (validation) spectra. The former group (244 spectra) was used for the establishment of the PLS statistical models. To validate the calibration, an independent set of 29 spectra were randomly selected from all sets representing the five countries. These 29 spectra contain 1 cheese from Austria (3 spectra), 1 cheese from Finland (3 spectra), 2 cheeses from Germany (6 spectra), 2 cheeses from France (6 spectra) and 3 cheeses from Switzerland (9 spectra) and 2 other spectra from Switzerland. As the successful establishment of the PLS model consists of modelling and validation procedures, the leave-one-out cross-validation

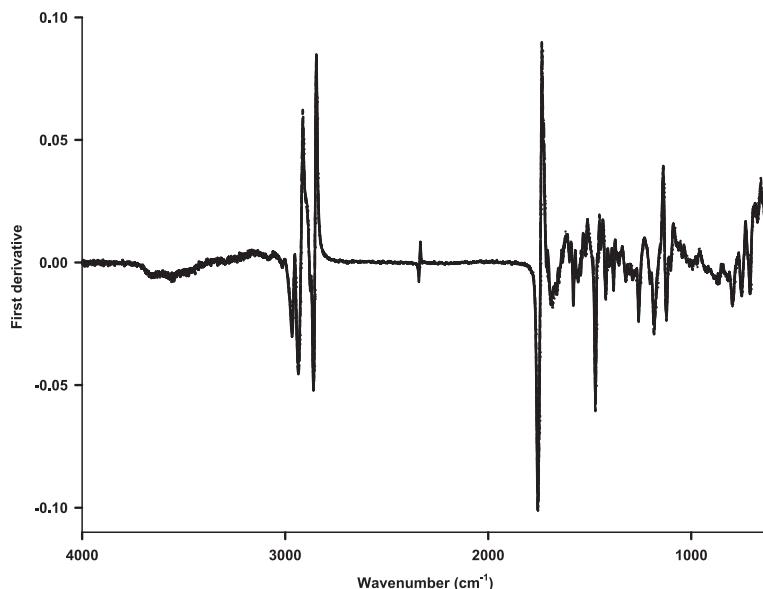
method was used according to the PLS modelling procedure. The first group that was designated as the calibration set of 244 (Tab. I) spectra was used for the combined PLS and cross-validation procedure. The established PLS models were further evaluated to predict different studied components using the remaining 29 spectra (Tab. II), which was designated as the validation set.

The practical utility of the calibrations was assessed using the range error ratio (RER) [39]. This ratio is calculated by dividing the range of a given constituent by the prediction error for that constituent and, while susceptible to the presence of extreme values at both ends of the range, it is a useful indicator of the practical utility of a predictive model.

### 3. RESULTS AND DISCUSSION

As each chemical compound in the cheese contributes to the absorbance spectrum, the MIR spectrum of a cheese contains information about the compounds present at levels > 0.1% (w/w). The absorption bands observed in the MIR range of 4000–600 cm<sup>-1</sup> are associated with fundamental valence vibrations of functional groups of the molecule. For a given cheese, a good repeatability is obtained from the three spectral measurements [19, 20].

In order to obtain a more detailed description of spectral variations, the first derivative of selected MIR spectra collected



**Figure 1.** Averaged first derivative of the MIR spectra for Emmental cheeses from Austria, Germany, Switzerland, France and Finland.

**Table II.** Sample statistics of validation set<sup>a</sup>.

Compositional parameter	Minimum	Maximum	Mean	Standard deviation
Fat (g·kg <sup>-1</sup> )	280.00	360.00	314.26	17.64
NaCl (g·kg <sup>-1</sup> )	2.50	13.60	4.96	2.52
pH	5.585	5.84	5.64	0.09
NPN (g·kg <sup>-1</sup> )	2.02	9.38	6.57	1.61
TN (g·kg <sup>-1</sup> )	40.14	46.89	44.39	1.40
WSN (g·kg <sup>-1</sup> )	3.74	13.16	9.35	2.18
WSN/TN	0.08	0.30	0.21	0.05
WSN/NPN	1.08	1.92	1.44	0.18

<sup>a</sup> Abbreviation: NPN, non-protein nitrogen; TN, total nitrogen; WSN, water-soluble nitrogen.

from different European regions is shown in Figure 1. Contribution to the lipids can be observed in the 3000–2800 cm<sup>-1</sup> spectral region, which is dominated by two strong bands at 2915 and 2846 cm<sup>-1</sup> associated with methylene anti-symmetric and sym-

metric stretching [9], respectively. Two other bands resulting from the asymmetric and symmetric stretching modes of the terminal methyl groups were also present at 2954 and 2860 cm<sup>-1</sup>, respectively. Contributions to amide I can be observed around

**Table III.** Correlation coefficients for physico-chemical variables analysed on Emmental cheeses<sup>a</sup>.

Variables	Fat	NaCl	pH	NPN	TN	WSN	WSN/ TN	WSN/ NPN
Fat	1							
NaCl	0.19*	1						
pH	0.24*	0.34*	1					
NPN	0.28*	0.08	0.28*	1				
TN	-0.39*	-0.04	0.15*	0.23*	1			
WSN	0.43*	0.24*	0.48*	0.84*	0.18*	1		
WSN/TN	0.50*	0.25*	0.47*	0.82*	0.03	0.99*	1	
WSN/NPN	0.18*	0.25*	0.24*	-0.36*	-0.07	0.18*	0.20*	1

<sup>a</sup> Abbreviation: NPN, non-protein nitrogen; TN, total nitrogen; WSN, water-soluble nitrogen; \* level significance  $P < 0.05$ .

1684 and 1622 cm<sup>-1</sup>. This part of the MIR was used by other researchers [26] to investigate the secondary structure of several proteins. The absorption bands at 1578, 1526 and 1512 cm<sup>-1</sup> are generally assigned to amide II vibrations. The absorption band around 1578 cm<sup>-1</sup> has been attributed to soluble carboxylic acids, such as lactate, which has a characteristic wavelength at 1575 cm<sup>-1</sup> as has been shown by Picque et al. [31]. The region 1500–600 cm<sup>-1</sup>, called the fingerprint region, refers to C-O and C-C stretching modes (1153–900 cm<sup>-1</sup>). Absorbance in the region 1474–1200 cm<sup>-1</sup> is due to bending modes of O-C-H, C-C-H and C-O-H [30, 37]. The strong absorption at 1470 cm<sup>-1</sup> is assigned to the carboxylic acid group of lactic acid [12]. The bands located between 1042 and 1000 cm<sup>-1</sup> have been attributed to lactate [31]. The P = O stretching band contributes strongly in the 1500–600 cm<sup>-1</sup>, as has been reported by Bellamy [3].

Principal component analysis (PCA) was used to investigate the inter-correlations between the investigated variables (fat, NaCl, pH, NPN, TN, WSN, WSN/TN and WSN/NPN) (Tab. III). High correlation between WSN and NPN ( $r = 0.84$ ,  $P < 0.05$ ), WSN/TN and NPN ( $r = 0.82$ ,  $P < 0.05$ ) and between WSN/TN and WSN ( $r = 0.99$ ,  $P < 0.05$ ) indicated that the ripening period affected both WSN and NPN in a similar way. Negative correlation was observed between protein and fat ( $r = -0.39$ ,  $P < 0.05$ ).

The PLS-cross-validation technique provided different calibration results according to the component being investigated. To compare between different models established for a given component, the values of the root mean square error of cross-validation (RMSECV) of the calibration set and prediction (RMSEP) for the prediction set are considered. The accuracy of each calibration can be evaluated based on the coefficient of determination ( $R^2$ ) for predicted versus measured compositions in cross-validation and prediction, and the ratio of standard deviation of the data set (SD) to RMSECV or RMSEP, called RPD. In fact,  $R^2$  indicates the percentage of the variance in the  $Y$  variable that is accounted for by the  $X$  variable. A value for  $R^2$  between 0.50 and 0.65 indicates that more than 50% of the variance in  $Y$  is accounted for by variable  $X$ , so that discrimination between high and low concentrations can be made. A value for  $R^2$  between 0.66 and 0.81 indicates approximate quantitative predictions, whereas a value for  $R^2$  between 0.82 and 0.90 reveals good prediction. Calibration models with a value for  $R^2$  above 0.91 are considered to be excellent [40]. The ratio of the SD to the RMSECV or the RMSEP, called the ratio of prediction to deviation (RPD), is the factor by which the prediction accuracy has been increased compared with using the mean composition for all samples. This ratio is

**Table IV.** Validation results of PLS-cross-validation regression on calibration sample set<sup>a</sup>.

Compositional parameter	LV	R <sup>2</sup>	Slope	RMSECV	RPD
Fat (g·kg <sup>-1</sup> )	12	0.69	0.73	7.28	1.87
NaCl (g·kg <sup>-1</sup> )	10	0.70	0.73	1.19	1.91
pH	11	0.82	0.84	0.04	2.34
NPN (g·kg <sup>-1</sup> )	5	0.85	0.85	0.54	2.58
TN (g·kg <sup>-1</sup> )	10	0.62	0.69	0.83	1.60
WSN (g·kg <sup>-1</sup> )	6	0.80	0.81	0.86	2.26
WSN/TN	10	0.83	0.84	0.02	2.81
WSN/NPN	10	0.57	0.61	0.12	1.52

<sup>a</sup> Abbreviation: NPN, non-protein nitrogen; TN, total nitrogen; WSN, water-soluble nitrogen; LV, latent variables; R<sup>2</sup>, determination coefficient; RMSECV, root mean square error of cross-validation; RPD, ratio of prediction to deviation (Standard deviation/RMSECV).

**Table V.** Validation of PLS-cross-validation regression on validation set<sup>a</sup>.

Compositional parameter	R <sup>2</sup>	Slope	RMSEP	RPD
Fat (g·kg <sup>-1</sup> )	0.57	0.54	8.33	1.55
NaCl (g·kg <sup>-1</sup> )	0.60	0.61	1.35	1.59
pH	0.84	0.85	0.03	2.55
NPN (g·kg <sup>-1</sup> )	0.83	0.74	0.54	2.34
TN (g·kg <sup>-1</sup> )	0.62	0.64	0.71	1.64
WSN (g·kg <sup>-1</sup> )	0.88	0.78	0.83	2.64
WSN/TN	0.90	0.83	0.02	2.94
WSN/NPN	0.64	0.70	0.11	1.70

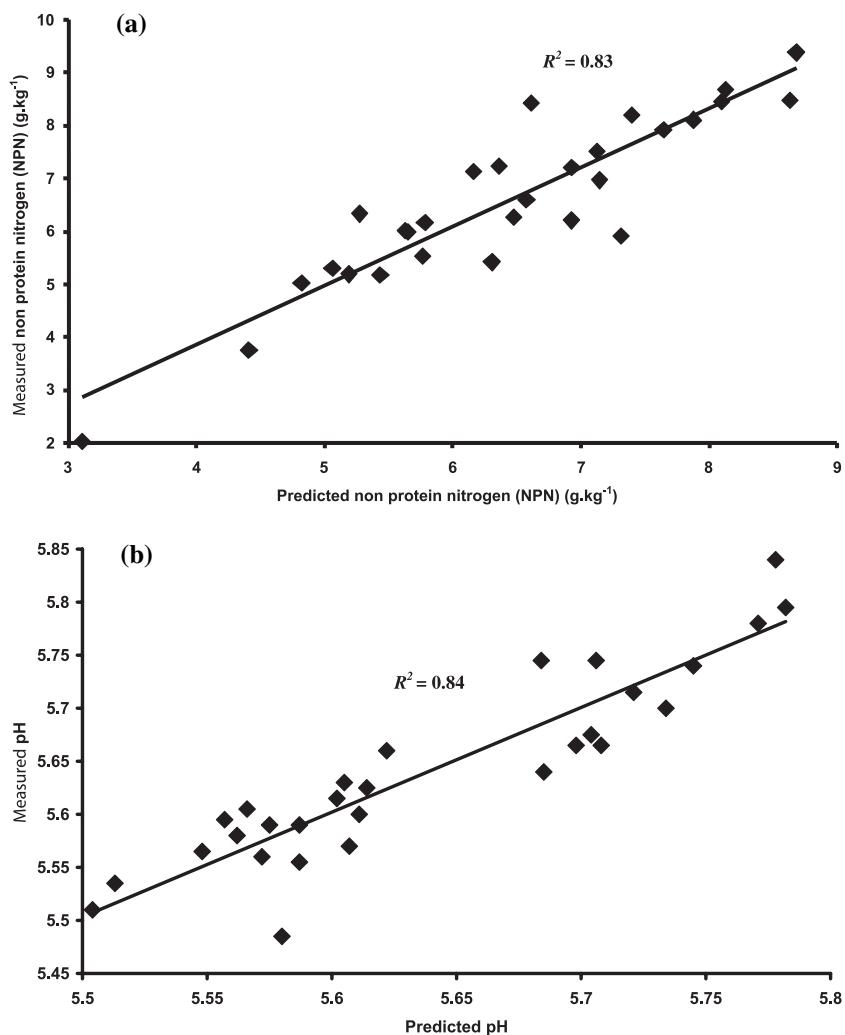
<sup>a</sup> Abbreviation: NPN, non-protein nitrogen; TN, total nitrogen; WSN, water-soluble nitrogen; R<sup>2</sup>, determination coefficient; RMSEP, root mean square error of prediction; RPD, ratio of prediction to deviation (Standard deviation/RMSEP).

desired to be larger than 2 for a good calibration [36]. A RPD ratio less than 1.5 indicates incorrect predictions and the model cannot be used for further prediction.

A summary of the prediction performance of the PLS models developed for each of the chemical parameters using the whole wavelength range is shown in Tables IV and V. The pH and NPN content were the best components to be predicted with the developed MIR-PLS methodology. According to the R<sup>2</sup> and RPD values, the prediction of these parameters can be evaluated as good, since the R<sup>2</sup> of NPN and pH were 0.85 and 0.82, respectively (Tab. IV). This was also

valid for the validation set (Figs. 2a and 2b), with results shown in Table V. The prediction of WSN could be considered as an approximate quantitative possible for the calibration data set and as good for the validation data set (R<sup>2</sup> = 0.88 and RPD = 2.64) (Tab. V). These results are confirmed by unpublished findings in our laboratories reporting that MIR could be used for the determination of the NPN and WSN of European Emmental cheeses produced during summertime.

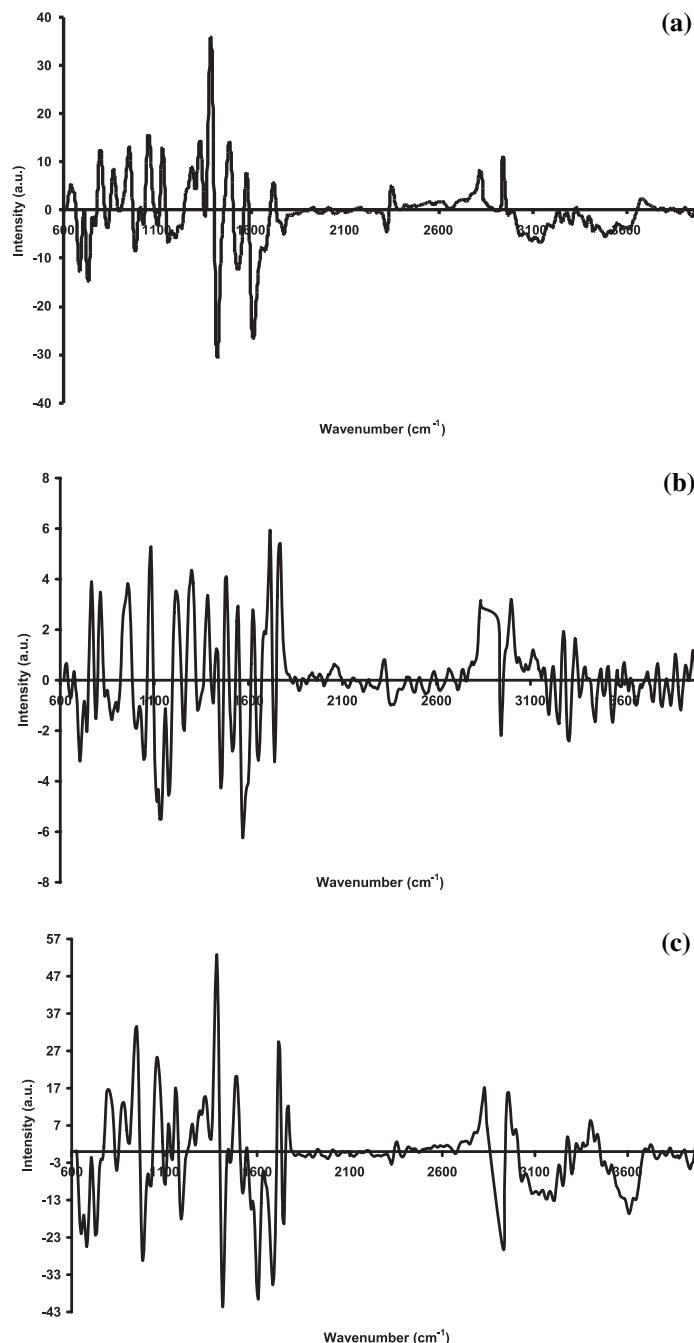
Regarding NPN content, the regression coefficient distribution over the entire wavelength range implies that the 1734–600 cm<sup>-1</sup>



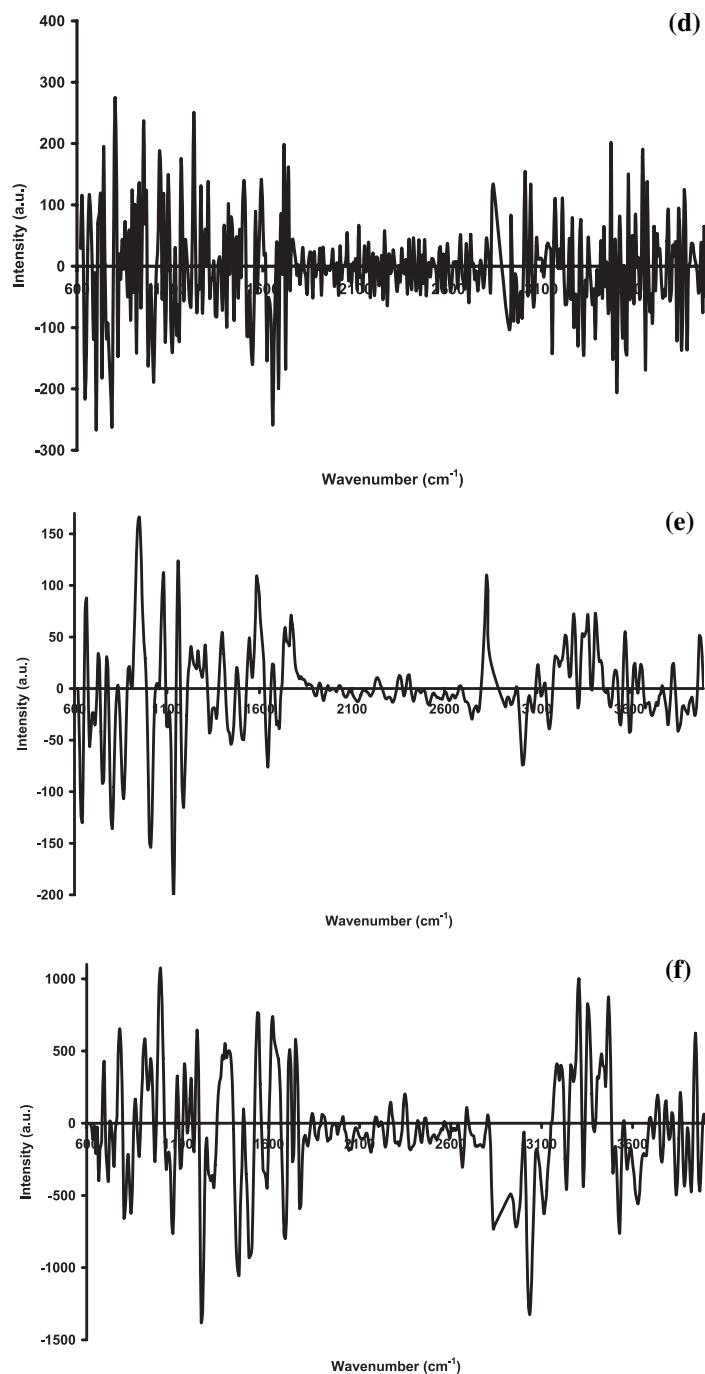
**Figure 2.** Linear regression plot of actual versus predicted non protein nitrogen content (NPN) (a) and pH (b) of the validation data set.

spectral region and the bands located at 2827 and 2934 cm<sup>-1</sup> are the most important regions due to the presence of several correlation peaks (Fig. 3a). The regression coefficient distribution for the pH shows that both the 1738–600 and the 4000–2834 cm<sup>-1</sup> spectral bands are the most efficient regions (Fig. 3b). While only the wavelength range of 1738–600 cm<sup>-1</sup> and the bands located at the 2827 and 2934 cm<sup>-1</sup> spectral

regions were considered during the PLS-cross-validation stage, the NPN and pH models provide very similar results to the corresponding models developed for the entire wavelength range (data not shown). This is in agreement with what is shown by the first derivative of the spectra (Fig. 1), where more variable absorbance can be seen in both regions of 1738–600 cm<sup>-1</sup> and 2968–2846 cm<sup>-1</sup>.



**Figure 3.** Regression coefficient distributions over the entire wavelength range of non protein nitrogen (NPN) (a), pH (b), water soluble nitrogen (WSN) (c), total nitrogen (TN) (d), NaCl (e) and fat (f).



**Figure 3.** Continued.

According to the  $R^2$  and RPD, the prediction of the level of proteolysis expressed by the WSN/TN and NPN/TN ratios could be considered as good for the WSN/TN ( $R^2 = 0.83$ ) and as possible to discriminate between high and low concentrations for the NPN/TN ( $R^2 = 0.57$ ). This was also valid for the validation set, with results shown in Table V. Although a low prediction performance was obtained for the NPN/TN using the MIR-PLS technique, the results obtained for WSN/TN are of economic importance for the cheese industry. Up-to-date cheese ripening is a food process that is not highly automated. Cheese ripening as a quality indicator of the final products depends on the know how and on the skills of the cheese maker. In regard to this concept, the work conducted in this study aimed at acquiring information about the ripening state of a hard cheese variety, which is needed for the cheese maker. The data collected enabled the development of models to help the cheese maker to estimate the level of cheese maturation during ripening. The level of proteolysis determined by the WSN/TN ratio brings satisfactory estimation of the ripening time required. In this way, the development of a fast methodology enabling the prediction of the quality of ripened cheese will allow the cheese-maker to act fast by anticipating or correcting a deviation that might reduce the price of the product due to quality deterioration. In addition, the cheese-maker can determine with good accuracy when the cheeses have reached maturation. Although the technique used in this study is destructive, an alternative and non-destructive technique such as FTIR-ATR could be used for the determination of the level of proteolysis.

An approximate quantitative prediction was observed for the NaCl and fat contents with  $R^2$  values of 0.70 and 0.69, respectively (Tab. IV). Less successful prediction of TN was obtained ( $R^2 = 0.62$ ) (Tab. IV). However, the obtained result of TN is in agreement with previous findings reporting that the correlation coefficient ( $r$ ) of the prediction for protein is 0.81 using FTIR-ATR [28]. Unpublished results in our laboratories found that MIR is not a suitable technique for the determination of TN and fat

contents in European Emmental cheeses produced during summer time. However, NIR could determine these two parameters with good accuracy. These results support previous findings reporting that better predictions of protein and fat contents were obtained with NIR than with MIR [28]. In addition to the simplicity of the NIR instruments compared with the MIR instruments, NIR is faster, more compact and robust [28]. Although the results obtained for TN are not good enough, those for WSN and NPN are of special interest because the Kjeldahl and trinitrobenzene-sulphonic acid (TNBS) methods of protein determination are very expensive and require skilled operators. Even though the TNBS method is rapid (approximately 2 h after the preparation of cheese samples), it is very expensive when looking at enzymatic and bio-chemical needs, which is not the case of MIR when the capital cost is excluded. The  $R^2$  and RMSEP values reported in this study for TN are in the same range as reported in previous investigations using NIR spectroscopy [32, 41]. However, the investigated Emmental cheeses herein have different sources of variability. They are collected from different European countries, manufactured from different milks (raw or thermised), using different cheese-making procedures and ranging in age from 12 to 16 weeks. Indeed, each region produces Emmental cheese with typical characteristics. Furthermore, compared with the above-mentioned study on using NIR spectroscopy, fewer latent variables (LV) (Tab. IV) were used in this study for NPN and WSN (5 and 6, respectively), which makes the models more robust.

Sample heterogeneity is already known to be critical for a satisfactory performance of MIR. To make progress in understanding the effect of sample heterogeneity on the measurement of MIR, a description of the heterogeneity should be made. In the case of cheese, heterogeneities at the micron scale level are expected primarily because of the fat globules [28]. In addition, it is known that fat and protein contents of large cheeses such as European Emmental cheese variety can vary by several percent according to the manufacture procedure (Tab. I).

Furthermore, for a given cheese, the region (feed type and weather conditions) where a sample is extracted from the cheese can have an effect on the MIR measurement. According to McQueen et al. [28], the only way to form an appropriate average for the physico-chemical parameters is to take samples from different parts of the cheese and then average the obtained results. The averaging procedure produces far more representative values than single measurements. Another explanation could arise from the water subtraction procedure, which allows for a significant reduction in interference due to water. This was the primary obstacle to the determination of protein concentration, as has been reported by Etzion et al. [11]. However, an accurate water subtraction sufficed for the determination of both protein and fat concentrations, and cheese spectra were still influenced by other constituents such as lactate, which affect protein-water and protein-fat interactions. These interactions could explain the inaccuracy of MIR transmission for the prediction of fat and TN.

The distribution of the regression coefficient of the WSN (Fig. 3c) shows similar results to those obtained for the NPN, while those for the NaCl (Fig. 3e) had a similar trend to those obtained for the pH. However, the distribution of the regression coefficient of the fat shows a rather different trend compared with the above-mentioned parameters (Fig. 3f). The 1777–600 cm<sup>-1</sup> and 3000–2800 cm<sup>-1</sup> spectral regions are the most important regions for fat. Considering these two spectral regions, the PLS-cross-validation gave similar results to those obtained using the region 4000–600 cm<sup>-1</sup> (data not shown). Although the distribution of the regression coefficient of TN (Fig. 3d) is too noisy, it is quite similar to those of WSN and NPN.

RMSECV for fat, NaCl, pH, NPN, TN, WSN, WSN/TN and WSN/NPN were 7.28, 1.19, 0.04, 0.54, 0.83, 0.86, 0.02 and 0.12, respectively. These RMSECV values resulted in corresponding RER values of 10.98, 9.83, 11.5, 14.44, 8.30, 11.08, 11 and 11.17, respectively, suggesting that the models may have practical utility for WSN, NPN, pH and WSN/TN. This implies that by

using the developed MIR-PLS methodology it is possible to provide quantitative measurement of the pH, NPN, WSN and WSN/TN with different degrees of accuracy. Quantitative measurement of the fat, TN, WSN/NPN and NaCl is not supported by the methodology and results provided in this study. The result obtained for TN and WSN/NPN is quite surprising, since a good correlation was found for the WSN, NPN and TN. For the moment, no scientific explanation for the unsuccessful prediction of TN and WSN/NPN from the MIR can be provided.

#### 4. CONCLUSION

The results obtained in this study clearly indicate the suitability of MIR spectroscopy for the determination of NPN, pH, WSN and the proteolysis level expressed by the WSN/TN ratio of European Emmental cheeses with different origins produced in winter. However, the results obtained for fat, NaCl, TN and WSN/NPN are not good enough, with RPD values smaller than 2 and  $R^2$  values of 0.69, 0.70, 0.62 and 0.57, respectively. More research is needed in order to test the ability of this technique for the prediction of TN, NaCl and fat. Furthermore, the accuracy of the method should be investigated on Emmental cheese produced during summertime, since there is a significant effect of the season on the characteristics of the investigated cheeses.

#### REFERENCES

- [1] Adamopoulos K.G., Goula A.M., Petropakis H.J., Quality control during processing of Feta cheese-NIR application, *J. Food Compos. Anal.* 14 (2001) 431–440.
- [2] Adams M.J., Latham K., Barnett N.W., Poynton A., Calibration models for determining moisture and fat content of processed cheese using near-infrared spectrometry, *J. Sci. Food Agric.* 79 (1999) 1232–1236.
- [3] Bellamy L.J., The infrared spectra of complex molecules, John Wiley and Sons, Inc., New-York, 1975.

- [4] Cattaneo T.M.P., Giardina C., Sinelli N., Riva M., Giangiacomo R., Application of FT-NIR and FT-IR spectroscopy to study the shelf-life of Crescenza cheese, *Int. Dairy J.* 15 (2005) 693–700.
- [5] Chen M., Irudayaraj J., McMahon D.J., Examination of full fat and reduced fat Cheddar cheese during ripening by Fourier infrared spectroscopy, *J. Dairy Sci.* 81 (1998) 2791–2797.
- [6] Collomb M., Spahni-Rey M., Steiger G., Dosage de la teneur en azote selon Kjeldahl de produits laitiers et de certaines de leurs fractions azotées à l'aide d'un système automatisé, *Trav. Chim. Aliment. Hyg.* 81 (1990) 499–509.
- [7] Curda L., Kukacková O., NIR spectroscopy: a useful tool for rapid monitoring of processed cheeses manufacture, *J. Food Eng.* 61 (2004) 557–560.
- [8] Desmazeaud M.J., Gripon J.C., General mechanism of protein breakdown during cheese ripening, *Milchwissenschaft* 32 (1977) 731–734.
- [9] Dufour E., Mazerolles G., Devaux M.F., Duboz G., Duployer M.H., Mouhous Riou N., Phase transition of triglycerides during semi-hard cheese ripening, *Int. Dairy J.* 10 (2000) 81–93.
- [10] Dufour E., Karoui R., Bosset J.O., Utilisation de la fluorescence frontale intrinsèque de fromages de type L'Etivaz AOC et Gruyère AOC pour reconnaître leur origine géographique, *Trav. Chim. Aliment Hyg.* 94 (2003) 379–393.
- [11] Etzion Y., Linker R., Cogan U., Shmulevich I., Determination of protein concentration in raw milk by mid-infrared Fourier Transform infrared/Attenuated total reflectance spectroscopy, *J. Dairy Sci.* 87 (2004) 2779–2788.
- [12] Fayolle P., Picque D., Corrieu G., Monitoring of fermentation processes producing lactic acid bacteria by mid-infrared spectroscopy, *Vib. Spectrosc.* 14 (1997) 247–252.
- [13] Herbert S., Caractérisation de la structure moléculaire et microscopique de fromages à pâte molle. Analyse multivariée des données structurales en relation avec la texture, Thèse, École Doctorale Chimie Biologie de l'Université de Nantes, France, 1999.
- [14] Herbert S., Mouhous Riou N., Devaux M.F., Riaublanc A., Bouchet B., Gallant J.D., Dufour E., Monitoring the identity and the structure of soft cheeses by fluorescence spectroscopy, *Lait* 80 (2000) 621–634.
- [15] Hermida M., Gonzalez J.M., Sanchez M., Rodríguez-Otero J.L., Moisture, solids-non-fat and fat analysis in butter by near infrared spectroscopy, *Int. Dairy J.* 11 (2001) 93–98.
- [16] IDF, Cheese and processed cheese products. Determination of chloride content. Potentiometric titration method, Standard 88A, Int. Dairy Fed. Brussels, Belgium, 1988.
- [17] ISO, Cheese determination of fat content. Van Gulik method. Standard 3433, Int. Standards Organ., Geneva, Switzerland, 1975.
- [18] Karoui R., Contribution à l'étude des propriétés rhéologiques et à la détermination de l'origine géographique des fromages aux moyens des méthodes spectroscopiques et chimiométriques, Thèse, École Doctorale de Clermont-Ferrand II, France, 2004.
- [19] Karoui R., Dufour E., Pillonel L., Picque D., Cattenoz T., Bosset J.O., Fluorescence and infrared spectroscopies: a tool for the determination of the geographic origin of Emmental cheeses manufactured during summer, *Lait* 84 (2004) 359–374.
- [20] Karoui R., Dufour E., Pillonel L., Picque D., Cattenoz T., Bosset J.O., Determining the geographic origin of Emmental cheeses produced during winter and summer using a technique based on the concatenation of MIR and fluorescence spectroscopic data, *Eur. Food Res. Technol.* 219 (2004) 184–189.
- [21] Karoui R., Bosset J.O., Mazerolles G., Kulmyrzaev A., Dufour E., Monitoring the geographic origin of both experimental French Jura hard cheeses and Swiss Gruyère and l'Etivaz PDO cheeses using mid-infrared and fluorescence spectroscopies, *Int. Dairy J.* 15 (2005) 275–286.
- [22] Karoui R., Dufour E., Pillonel L., Schaller E., Picque D., Cattenoz T., Bosset J.O., The potential of combined infrared and fluorescence spectroscopies as a method of determination of the geographic origin of Emmental cheeses, *Int. Dairy J.* 15 (2005) 287–298.
- [23] Laporte M.F., Martel R., Paquin P., The near-infrared optic probe for monitoring rennet coagulation in cow's milk, *Int. Dairy J.* 8 (1998) 659–666.
- [24] Lebecque A., Laguet A., Devaux M.F., Dufour E., Delineation of the texture of Salers cheese by sensory analysis and physical methods, *Lait* 81 (2001) 609–623.
- [25] Mazerolles G., Duboz G., Hugot S., Détermination du taux d'humidité et de matière grasse de fromages de type pâte pressée par spectroscopie proche infrarouge en mode transmission, *Lait* 80 (2000) 371–379.
- [26] Mazerolles G., Devaux M.F., Duboz G., Duployer M.H., Mouhous Riou N., Dufour E., Infrared and fluorescence spectroscopy for monitoring protein structure and interaction changes during cheese ripening, *Lait* 81 (2001) 509–527.

- [27] Mazerolles G., Devaux M.F., Dufour E., Qannari E.M., Courcoux P., Chemometric methods for the coupling of spectroscopic techniques and for the extraction of the relevant information contained in the spectral data tables, *Chemometr. Intell. Lab.* 63 (2002) 57–68.
- [28] Mc Queen D.H., Wilson R., Kinnunen A., Jensen E.P., Comparison of two infrared spectroscopic methods for cheese analysis, *Talanta* 42 (1995) 2007–2015.
- [29] O'Callaghan D.J., O'Donnell C.P., Payne F.A., On-line sensing techniques for coagulum setting in renneted milks, *Int. Dairy J.* 10 (2000) 155–165.
- [30] Paradkar M.M., Sivakesava S., Irudayaraj J., Discrimination and classification of adulterants in maple syrup with the use of infrared spectroscopic techniques, *J. Sci. Food Agric.* 83 (2003) 714–721.
- [31] Picque D., Lefier D., Grappin R., Corrieu G., Monitoring of fermentation by infrared spectrometry—Alcoholic and lactic fermentations, *Anal. Chim. Acta* 279 (1993) 67–72.
- [32] Pierce M.M., Wehling R.L., Comparison of sample handling and data treatment methods for determining moisture and fat in Cheddar cheese by near infrared spectroscopy, *J. Agric. Food Chem.* 42 (1994) 2831–2835.
- [33] Pillonel L., Luginbühl W., Picque D., Schaller E., Tabacchi R., Bosset J.O., Analytical methods for the determination of the geographic origin of Emmental cheese. Mid and near-infrared spectroscopy, *Eur. Food Res. Technol.* 216 (2003) 179–183.
- [34] Pillonel L., Badertscher R., Casey M., Meyer J., Rossmann A., Schlichtherle-Cerny H., Tabacchi R., Bosset J.O., Geographic origin of European Emmental cheese: characterisation and descriptive statistics, *Int. Dairy J.* 15 (2005) 547–556.
- [35] Rodriguez-Otero J.L., Hermida M., Centeno J., Analysis of dairy products by near-infrared spectroscopy: A review, *J. Agric. Food Chem.* 45 (1997) 2815–2819.
- [36] Sinnaeve G., Dardenne P., Agneessens R., Biston R., The use of near infrared spectroscopy for the analysis of fresh grass silage, *J. Near Infrared Spectrosc.* 2 (1994) 79–84.
- [37] Sivakesava S., Irudayaraj J., A rapid spectroscopic technique for determining honey adulteration with corn syrup, *J. Food Sci.* 66 (2001) 787–792.
- [38] Sørensen L.K., Jepsen R., Assessment of sensory properties of cheese by near infrared spectroscopy, *Int. Dairy J.* 8 (1998) 863–871.
- [39] Williams P.C., Variables affecting near infrared reflectance spectroscopic analysis, in: Williams P., Norris K. (Eds.), *Near infrared technology in the agricultural and food industries*, AACC, St Paul, USA, 1987, pp. 143–167.
- [40] Williams P., *Near-Infrared Technology – Getting the Best out of Light*, PDK Grain, Nanaimo, Canada, 2003.
- [41] Wittrup C., Nørgaard L., Rapid near infrared spectroscopic screening of chemical parameters in semi-hard cheese using chemometrics, *J. Dairy Sci.* 81 (1998) 1803–1809.