Interpretative summary: Visible and near-infrared bulk optical properties of raw milk. Aernouts.

The quality of milk is important for the dairy farmer, milk processing plants, retail and the consumer. Optical techniques based on Vis/NIR spectroscopy have already proven their potential for automated monitoring of the milk composition and microstructure as these properties are related to respectively the absorption and scattering of light. Nevertheless, the interaction between absorption and scattering of the light travelling through the sample complicates the interpretation of the measured signals. Therefore, the sensor should be well designed and combined with a robust light propagation model to obtain accurate predictions of the milk properties. In this paper, the Vis/NIR bulk optical properties of raw milk are studied and reported. This information is essential for the optimization of a Vis/NIR optical milk quality sensor.

VISIBLE AND NEAR-INFRARED BULK OPTICAL PROPERTIES OF RAW MILK

Visible and near-infrared bulk optical properties of raw milk

B. Aernouts, R. Van Beers, R. Watté, T. Huybrechts, J. Lammertyn, and W. Saeys

KU Leuven, Department of Biosystems, MeBioS, Kasteelpark Arenberg 30, 3001 Leuven, Belgium.

Corresponding author: Ben Aernouts, KU Leuven, Department of Biosystems, MeBioS, Kasteelpark Arenberg 30, 3001 Leuven, Belgium. Phone: +32 16 321470, Fax: +32 16 321994, E-mail: Ben.Aernouts@kuleuven.be

ABSTRACT

The implementation of optical sensor technology to monitor the milk quality on dairy farms and milk processing plants would support the early detection of altering production processes. Basic visible and near-infrared (Vis/NIR) spectroscopy is already widely used to measure the composition of agricultural and food products. However, to obtain maximal performance, the design of such optical sensors should be optimized with regards to the optical properties of the samples to be measured. Therefore, the aim of this study was to determine the Vis/NIR bulk absorption coefficient, bulk scattering coefficient and scattering anisotropy spectra for a diverse set of raw milk samples originating from individual cow milkings, representing the milk variability present on dairy farms. Accordingly, this database of bulk optical properties can be used in future simulation studies to efficiently optimize and validate the design of an optical milk quality sensor. In a next step of the current study, the relation between the obtained bulk optical properties and milk quality properties was analyzed in detail. The bulk absorption coefficient spectra were found to mainly contain information on the water, fat and casein content, while the bulk scattering coefficient spectra were found to be primarily influenced by the quantity and the size of the fat globules. Moreover, a strong positive correlation ($R \geq 0.975$) was found between the fat content in raw milk and the measured bulk scattering coefficients in the 1300 – 1400 nm wavelength range. Relative to the bulk scattering coefficient, the variability on the scattering anisotropy factor was found to be limited. This is because the milk scattering anisotropy is nearly independent of the fat globule and casein micelle quantity, while it is mainly determined by the size of the fat globules. As this study shows high correlations between the sample's bulk optical properties and the milk composition and fat globule size, a sensor which allows for robust separation between the absorption and scattering properties would enable accurate prediction of the raw milk quality parameters.

Keywords: Milk, fat globule, casein micelle, Visible and near-infrared spectroscopy, scattering, absorption, optical sensor design

INTRODUCTION

A precondition for increased profitability in dairy farming is an increase in both the lactation and lifetime production per cow. Therefore, more effective prevention and early treatment of all diseases, especially the so-called 'production diseases', is needed (Hamann and Krömker, 1997). To meet these demands, individual cow and udder health should be carefully monitored. Since the milk production is a dominant factor in the metabolism of dairy cows, involving a very intensive interaction with the blood circulation, the extracted milk contains valuable information on the nutritional, metabolic and infectious status of the cow (Aernouts et al., 2011; Løvendahl et al., 2010; Forsbäck et al., 2010, 2009; Friggens et al., 2007; Mulligan et al., 2006; Hamann and Krömker, 1997). Therefore, regular analysis of the produced milk is considered to be the most efficient way to monitor cow and udder health. Online measurement of the milk components (fat, protein, lactose, etc.) during milking twice a day would offer the potential for early detection of systemic and local alteration, thus providing a valuable input for strategic and operational management decisions (Friggens et al., 2007).

Visible (Vis) and near-infrared (NIR) spectroscopic analysis of raw milk allows for a reliable detection of the fat, protein and lactose concentration in the lab (Aernouts et al., 2011). The prediction of this milk composition is mainly based on the wavelength-dependent absorption of Vis/NIR radiation by the milk constituents. The industry has recently adopted this technology, and implemented it into milking systems to measure the major milk components on-line (Pinsky et al., 2013; Katz et al., 2011, 2003). However, despite of the continuous recalibration, their accuracy and robustness is still not sufficient to support cow health management (Kaniyamattam and De Vries, 2014). This is mainly because the measured spectral signals are, next to absorption, considerably influenced by the physical properties of the milk

in terms of the quantity and size of the fat globules and casein micelles. Since the refractive indices of milk fat and casein differ from that of the milk serum, the Vis/NIR radiation is forced to deviate from its straight trajectory (Tuchin, 2007). Because of these scattering processes, the travelling path of the radiation increases to an unknown extend. This complicates the prediction of the composition from measured spectra. Homogenization of the milk fat globules could reduce and standardize the Vis/NIR scattering to improve the prediction results. For example, in the early days of milk analysis, the Milko-tester (Foss Electric, Hillerød, Denmark) measured the Vis scattering after dispersion of the casein micelles and homogenization to produce a more uniform fat globule size distribution. Accordingly, the light attenuation depends mainly on the amount of fat globules and can, therefore, be used to obtain a rough estimate for the fat content (McDowell, 1968). However, because of its destructive character, high energy consumption and significant wear and tear, the proceeding homogenization step is not desired in online analyses on-farm. Alternatively, the non-linear interference due to light scattering can be reduced with empirical methods (e.g. baseline correction, derivatives, …) or over-simplistic scattering models (e.g. normal, piecewise and extended multiplicative scatter correction, path length correction method, …) and/or can be partially accounted for by the prediction models (e.g. partial least squares, support vector machines, …) (Aernouts et al., 2011). However, these techniques only provide acceptable results for samples with absorption and scattering properties similar to those consulted in the calibration procedure (Melfsen et al., 2013). As the quantity and size of the fat globules and casein micelles for different fresh raw milk samples experiences large variations, the scattering properties too vary a lot (Logan et al., 2014; Cabassi et al., 2013; Nielsen et al., 2005; Vangroenweghe et al., 2002). Consequently, it is very challenging to fully compensate for all this scattering variability with a single empirical calibration model. Therefore, more powerful and advanced techniques are needed to remove the scattering interference from the measured Vis/NIR spectra (Melfsen et al., 2012).

In Vis/NIR spectroscopy, accurate separation of the absorption and scattering properties would reduce the need for empirical scatter corrections and promote robust prediction of the sample composition (Steponavičius and Thennadil, 2013, 2011; Steponavicius and Thennadil, 2009). Moreover, the pure absorption, defined as the bulk absorption coefficient μ_a (cm^{-1}), is the probability of absorption per unit infinitesimal path length at a specific radiation wavelength and relates directly to the sample composition according to the Beer-Lambert law. The scattering, on the other hand, can be described with the bulk scattering coefficient μ_s (cm^{-1}) and the angular scattering pattern or scattering phase function. The bulk scattering coefficient defines the probability of scattering per unit infinitesimal path length in a similar way as μ_a represents the absorption. The scattering phase function is generally too complex to reproduce and interpret and is, therefore, often represented by its mean cosine: the scattering anisotropy factor *g*. The scattering anisotropy for biological tissues and fluids in the Vis/NIR range varies between 0 (isotropic scattering) and 1 (complete forward scattering) (Tuchin, 2007). These scattering properties are determined by the physical microstructure properties of the sample (e.g. particle size distribution, particle volume concentration, material properties, …). For milk, this primarily relates to the quantity and size of the suspended fat globules and, to a smaller extend, the casein micelles (Aernouts et al., 2015; Bogomolov et al., 2013; Bogomolov and Melenteva, 2013; Dahm, 2013; Kucheryavskiy et al., 2014; Bogomolov et al., 2012). As these properties affect the physicochemical, functional and sensory characteristics of the raw milk and derived dairy products, they are important quality parameters (Cabassi et al., 2013; Schenkel et al., 2013; Walstra et al., 2006; Michalski et al., 2004, 2003). Moreover, the size of fat globules in milk from infected udder quarters (mastitis) is increased significantly and could, therefore, give insight into the udder health status of each individual cow and udder quarter (Mizuno et al., 2012; Erwin and Randolph, 1975). Accordingly, extraction of physical microstructure information, such as the fat globule size distribution, from isolated scattering

properties would create an added value for Vis/NIR spectroscopy on raw milk (Aernouts et al., 2015; Cabassi et al., 2013).

In a single Vis/NIR spectroscopic measurement, usually reflectance or transmittance, both the effect of absorption by the chemical molecules and scattering by the physical particles are inter-connected and cannot be accurately separated. Consequently, a change in the scattering properties of a measured milk sample might be misinterpreted as a change in the milk composition (Melfsen et al., 2012). On the other hand, multiple spectroscopic measurements in a slightly different configuration are not perfectly correlated and will, therefore, be influenced by absorption and scattering in a different way. The combination of such multiple measurement series with an accurate model, which mathematically describes light propagation as a function of the sample's bulk optical properties (μ_a, μ_s and g), could provide a successful separation of the sample's absorption and scattering properties (Steponavičius and Thennadil, 2013). However, superior separation between these absorption and scattering properties is only feasible if the optical sensor is designed to obtain a series of multiple measurements with least inter-correlation and maximum signal-to-noise levels. As the measured signals are, next to the sensor architecture, determined by the sample's bulk optical properties (BOP), the optimal design of such a practical sensor configuration depends on the absorption and scattering properties of the samples to be measured. The effect of the BOP on the light propagation, and consequently the collected spectral signals, is very complex. Therefore, the optimal design cannot be calculated directly from a supplied range of BOP, though it can be determined through an iterative optimization procedure. In practice, a wide range of sensor configurations is physically possible. So, it is preferred to test the potential of each sensor configuration through these simulations, rather than building each of them and evaluating their performance from measurements on an extensive set of raw milk samples (Zamora-Rojas et al., 2014; Khankin et al., 2012; Gamm et al., 2011; Cen et al., 2010; Palmer and Ramanujam, 2007;

Sharma et al., 2006; Liu and Ramanujam, 2006; Luo et al., 2005). The sensor configuration which allows for the most robust separation between the absorption and scattering properties would obviously have the highest potential to retrieve accurate predictions for the milk composition (fat, protein, lactose, urea, etc.) and physical properties (fat globule and casein micelle size distribution) from respectively the obtained absorption and scattering properties. As these milk quality properties are highly correlated to cow health, such sensor would support on-farm dairy management.

The Monte Carlo (MC) method for simulation of light propagation is a very accurate, flexible and is widely used in tissue optics (Tuchin, 2007). Therefore, it is particularly suitable to simulate the collected spectra series for each sensor configuration and for the range of BOP found in raw milk samples. Accordingly, the potential of each sensor configuration can be defined as the ability to extract the BOP from the collected signals, after adding noise typical for a Vis/NIR spectrometer. Such procedures of sensor design optimization have been widely studied and improved in the last decade and are still an important topic of research (Zamora-Rojas et al., 2014; Khankin et al., 2012; Gamm et al., 2011; Cen et al., 2010; Palmer and Ramanujam, 2007; Sharma et al., 2006; Liu and Ramanujam, 2006; Luo et al., 2005). However, in order to consult these algorithms to obtain an optimal sensor design for quality control of raw milk, knowledge on the Vis/NIR BOP of raw milk is crucial.

Recently, the influence of a varying fat globules size on the Vis/NIR scattering properties of milk was studied in detail (Aernouts et al., 2015). Moreover, reduction of the fat globule size resulted in a higher wavelength-dependency of both the bulk scattering coefficient and the scattering anisotropy factor, reducing their values for wavelengths above 600 nm and approaching the Rayleigh scattering phenomenon. Nevertheless, to our knowledge, no accurate information is available in literature on the mean, the variability and the range of the Vis/NIR bulk optical properties of raw milk from dairy cows, therefore defined as the main objective of

this study. The measurements focus on the 550 – 1900 nm wavelength range as, below 550 nm, the main milk components have no relevant absorption peaks, while above 1900 nm, water is a very strong absorber resulting in very low signal-to-noise levels for any type of optical measurement. Next, the obtained data was used to closely study the relation between the Vis/NIR bulk optical properties and the milk's chemical and physical quality properties.

MATERIALS AND METHODS

Milk samples

The milk samples considered in this research were collected in the context of the milk production registration system within Flanders (Belgium). Dairy farmers from all over Flanders can participate in this system to monitor the milk composition and production of their individual cows ones every 4 – 6 weeks. These data are used to improve breeding and genetic selection and to some extent for the evaluation and basic adjustment of feeding. For each cow, a representative milk sample (27 ml) is collected, preserved (4°C and ±0.11% *v*/*v* preservative; Qlip N.V., Leusden, the Netherlands) and analyzed with the Milkoscan FT+ (Foss A/S, Hillerod, Denmark) to determine the milk fat and crude protein content (ISO 9622:2000). According to the fabricant, the preservative contained Patent Blue V calcium salt (CAS: 3536-49-0, Sigma-Aldrich, St. Louis, MO) as a visible colorant marker. For this study, 60 raw milk samples, originating from 60 different cows and 17 different dairy farms were selected from a large collection of 1200 samples (20 dairy farms) to cover the maximum range of the compositional variance. Moreover, the 1200 samples were ordered on fat content and the first and every 40^{th} sample was selected (total 31 samples). The same procedure was repeated for the protein content on the remaining samples.

Table 1 gives an overview of the most important statistical parameters describing the fat and crude protein content of the selected sample set. The casein content was calculated as 75.5% of the crude protein content (Aernouts et al., 2015; Walstra et al., 2006). Comparison of the

mean, standard deviation (SD) and the range (Max – Min) of this sample set [Table 1] with the same statistical parameters of a much larger dataset (Milk Control Center-Flanders) indicates that the 60 samples are representative for the large population of milk produced by individual cows in Flanders (Aernouts et al., 2011). The correlation between both components was found to be in the normal range (R = 0.34) (Aernouts et al., 2011).

Measurement of bulk optical properties for raw milk samples

Double integrating sphere (DIS) and unscattered transmittance measurements were used to determine the BOP of the milk samples, as this is considered to be the 'golden standard' method for BOP measurement of thin samples of turbid media. The sample illumination in this setup was especially designed to obtain high signal-to-noise spectra in the 500 – 2250 nm wavelength range for very turbid media like raw milk. It consists of a supercontinuum laser light source (500 – 2250 nm, 4 Watt optical power) in combination with a high-precision monochromator. The total reflectance (M_R) and total transmittance (M_T) were measured simultaneously on each milk sample loaded in a cuvette (Schott, Germany) with a path length of 600 µm and positioned between the two integrating spheres. Both spheres were equipped with a Vis (400 – 1100 nm) and NIR (1100 – 2400 nm) detector. Unscattered transmittance (M_U) was measured in a separate path with the Vis and NIR detectors positioned 1.5 m behind the sample to limit the fraction of scattered photons collected by the detectors (Aernouts et al., 2014, 2013). To obtain sufficient unscattered transmittance signal, the sample was loaded in a thinner cuvette with a path length of 170 µm (Schott, Germany). A series of slits between sample and detector further reduced the number of scattered photons captured in the unscattered transmittance signal. For a more extensive description of the measurement setup, the calibration and measurement procedure and a thorough validation, the reader is referred to (Aernouts et al., 2013). Moreover, this validation study showed the high repeatability and signal-noise ratio of

225 the system to obtain the BOP of very turbid samples in the Vis/NIR (Aernouts et al., 2013). The
226 samples were thoroughly stirred before they were measured at 22±1°C (room temperature) to
227 ensure the homogeneity and temperature stability of the sample during the measurement. All
228 sample spectra were measured from 550 until 1900 nm in steps of 10 nm by automated scanning
229 of the pre-dispersive monochromator. The measurement takes 110 seconds, which was well
230 below the time span (10 minutes) after which creaming was starting to have a measurable effect
231 on the collected signals.

232 The diffuse reflectance (M_R) of the samples was derived from the total reflectance after
233 subtraction of the specular reflectance. The latter was calculated at the air-cuvette and cuvette-
234 sample interfaces through the Fresnel equations which use the real refractive indices of air (1),
235 the cuvette windows (provided by the manufacturer, Schott, Germany), and the milk sample.
236 The refractive index was calculated for each sample individually from the available milk
237 composition data, with the equation proposed by Walstra and Jenness (1984), taking into
238 account the sample temperature (22°C).

239 The inverse adding doubling (IAD) routine developed and optimized by Prahl (Prahl,
240 2010) was consulted to obtain the Vis/NIR BOP spectra from the obtained diffuse reflectance
241 and total and unscattered transmittance spectra. Because of significant contribution of scattered
242 photons, no BOP estimation could be established if the unscattered transmittance was below
243 0.01%. This was the case for approximately one third of the samples, mainly for radiation
244 wavelengths shorter than 1200 nm. If M_U was below 0.01%, this measurement was neglected
245 and an estimate for the anisotropy factor g was provided to the IAD algorithm to allow for the
246 separation of μ_a and μ_s (Prahl, 2010). For these samples, the average g spectrum was used as
247 estimate. Moreover, as the variability between the obtained g spectra was very small [Figure
248 1(c)], the average g spectrum is expected to be close to the actual g spectrum and the separation
249 between scattering and absorption should be sufficiently accurate (Prahl, 2010). Additionally,

also the reduced scattering coefficient μ_s' is reported. μ_s' combines μ_s and g according to the similarity relation $\mu_s' = \mu_s(1 - g)$ and can be used to accurately describe scattering after sufficient scattering events. In other words, after diffusion of the light, scattering can be accurately described with μ_s' alone, without the need for separation between μ_s and g (Tuchin, 2007).

RESULTS AND DISCUSSION

Variability in the bulk optical properties of raw milk

The BOP spectra for all 60 raw milk samples were extracted from the measured M_R, M_T and M_U spectra with the IAD algorithm. In Figure 1, the mean, mean ± standard deviation (SD), minimum (Min) and maximum (Max) values are shown for the derived BOP. The μ_a spectra [Figure 1(a)] indicate a very clear signature of water, with absorption peaks around 970, 1200, 1450 and 1940 nm. It should be noted that the peak in μ_a around 650 nm is caused by the colorant (Patent Blue V), present in the added preservative. Most of the variation in the μ_a spectra can be noticed at the absorption peaks of the colorant, water and around 1220 and 1740 – 1770 nm. The latter wavelengths are typical absorption peaks for milk fat as they correspond to respectively the second and first overtone stretch-vibrations of the CH_2-bonds (Šašić and Ozaki, 2000). As the milk fat content varies between 1.52 and 12.0% (*v*/*v*) [Table 1], noticeable variation can be expected at those absorption bands. Moreover, because of the water displacement effect, a higher dry matter content, related to a higher fat and/or crude protein content, would result in a lower absorption at the water peaks, explaining the considerable variation at the water absorption peaks. Additionally, the high variability around 650 nm indicates that the preservative concentration clearly varies between samples. At wavelengths where nearly no absorption is expected (720 – 820 nm), still a small baseline of maximum 0.116 cm^{-1} can be noticed. This is probably the result of very little cross-talk between μ_a and μ_s in the BOP estimation procedure. As μ_s is relatively high (100 – 1000 cm^{-1}) compared to μ_a (0 – 35

cm^{-1}), little cross-talk of μ_s to μ_a is already noticeable as a small baseline, especially at wavelengths where μ_a is close to zero.

The variation in the scattering coefficient spectra of raw milk in the 550 – 1900 nm wavelength range is large, ranging from 100 until nearly 1000 cm^{-1} [Figure 1(b)]. Fat globules, and to a smaller extend also casein micelles, are the main cause of scattering in milk (Aernouts et al., 2015). As a result, a higher fat and casein (~protein) content, which is associated with a higher quantity of respectively fat globules and casein micelles, results in an increase of the bulk scattering coefficient spectra (Aernouts et al., 2014). Moreover, if scattering events are uncorrelated and the size of fat globules and casein micelles is stable, there is a linear positive (independent scattering) relation between μ_s and particle quantity (Aernouts et al., 2014; Gaygadzhiev et al., 2008; Alexander et al., 2002). As the variation in fat and protein content in the set of 60 samples is large, it is expected to be the main source of variation in the μ_s spectra. Additionally, as was found in an earlier study, also the variability in the fat globule size distribution between milk samples will have an important share in the μ_s variability, especially for wavelengths from 550 until 1100 nm (Aernouts et al., 2015). Moreover, smaller fat globules were found to result in a more steep Vis/NIR μ_s spectrum, with the maximum shifted towards smaller radiation wavelengths and vice versa (Aernouts et al., 2015). The μ_s variation in the set of 60 samples was found to be maximal in the 550 – 1100 nm wavelength range, with values ranging from 120 until 950 cm^{-1}. This is probably because the effect of the fat globule size on μ_s is maximal in this wavelength range, additional to the effect of the fat and protein content.

If the independent scattering condition is valid, the anisotropy spectrum should be mainly influenced by the size of the fat globules, while being independent of the fat content itself (Aernouts et al., 2014, 2015). As a result, the variability in the *g* spectra is relatively small [Figure 1(c)]. In the 550 – 1900 nm wavelength range, the anisotropy factor for raw milk increases steadily with increasing radiation wavelength until it reaches a maximum around 1000

nm (Aernouts et al., 2015). Around these wavelengths, the scattering anisotropy is maximal and the fat globules in raw milk scatter most of the light in the forward direction. For longer radiation wavelengths, the anisotropy factor decreases with increasing wavelength, indicating more isotropic scattering. In a previous study, it was found that a higher *g* spectrum, mainly for radiation wavelengths above 1100 nm, indicates larger milk fat globules (Aernouts et al., 2015).

As the reduced scattering coefficient spectrum is the result of both the μ_s and *g* spectrum, it contains information from both the milk fat globules size and quantity. However, as all the μ_s' spectra are nearly parallel [Figure 1(d)], it seems that the effects of the fat globule size on μ_s and *g* neutralize each other if they are combined. In the Vis/NIR, μ_s' follows a steady decrease with increasing radiation wavelength until it reaches a nearly stable level for wavelengths above 1500 nm.

Effect of the fat globules on the bulk scattering properties of raw milk

In earlier studies (Aernouts et al., 2015; Frisvad et al., 2007), it was shown that the fat globules are, next to the casein micelles, the main source of Vis/NIR scattering in unskimmed milk. This is because the volume fraction of the fat globules is usually larger, while the Vis/NIR scattering intensities for a normalized volume fraction are also higher (Aernouts et al., 2015). Additionally, relative to the fat content, the crude protein, which consists of ±75.5% *w*/*w* casein, experiences only small variations in individual raw milk samples [Table 1] (Aernouts et al., 2015; Walstra et al., 2006). As a result, the variability in the bulk scattering properties of raw unskimmed milk is mainly determined by the variability in the size and quantity of the fat globules. Because of this, the relation between the fat globule size and quantity, and the bulk scattering properties of raw milk is discussed more in detail. In Figure 2 the bulk scattering properties are shown for 6 raw milk samples with a varying fat content and a practically constant crude protein content (2.36 – 2.49 % *v*/*v*). The crude protein content was kept constant to further

reduce the effect of casein micelles on the interpreted results. In this plot, 3 groups of each 2 samples can be distinguished based on the fat content, with a large variability between groups and practically no variability within a group. This allows to study the effect of the fat content and fat globule size separately. The μ_s spectra [Figure 2(a)] indicate that a higher fat content generally results in higher Vis/NIR bulk scattering coefficients. However, within a group of similar fat content, large variability can still be noticed in the μ_s spectra, especially for wavelengths below 1100 nm. This is probably caused by a difference in size of the fat globules between the two samples in the same group. Within the low-fat group (3.45 ± 0.065% *v*/*v*), small differences between the μ_s spectra can only be noticed for the wavelengths below 1000 nm. The sample with 3.51% (*v*/*v*) fat probably contains slightly smaller fat globules, as an earlier study showed that smaller milk fat globules are related to a steeper μ_s spectrum in the 550 – 1900 nm range (Aernouts et al., 2015). The same phenomenon is even more clear for the other two groups. Additionally, in these groups, a steeper μ_s spectrum in the Vis/NIR range is also related to a maximum μ_s at smaller wavelengths, typical for smaller scattering particles (Aernouts et al., 2015; Cabassi et al., 2013; Cattaneo et al., 2009). Within each group of raw milk samples with a similar fat content, the μ_s spectra seem to cross each other in the 1200 – 1400 nm wavelength range. Accordingly, the μ_s in this wavelength region might be less dependent on the fat globule size and have a higher correlation with the fat content itself.

As the anisotropy factor should be independent of the fat content if scattering processes are independent, g mainly contains information on the size of the fat globules (Aernouts et al., 2015, 2014). Moreover, an earlier study indicated that larger milk fat globules resulted in a higher anisotropy factor in the 1100 – 1900 nm wavelength range (Aernouts et al., 2015). Within each group, the sample with the smallest fat globules, according to the μ_s spectra (steeper and maximum shifted towards smaller wavelengths), was also characterized by a lower g spectrum for wavelengths above 1100 nm. Moreover, even between groups, a lower g

spectrum in the 1100 – 1900 nm wavelength range strongly correlates with a steeper μ_s spectrum and a maximum μ_s at shorter wavelengths. This strengthens the hypotheses that were generated in the previous paragraph.

The effect of the fat globule size, which is unambiguously present in the μ_s and g spectra, is not clearly noticeable in the μ_s' spectra. Moreover, while the 3 fat-content groups could clearly be separated based on their μ_s spectra, only 2 distinct groups appear in the μ_s' spectra for wavelengths below 1200 nm. Furthermore, no grouping of the μ_s' spectra can be noticed for wavelengths above 1200 nm. So, it seems that the fat globule size information present in μ_s and g, and the fat content information in μ_s partially neutralize each other if μ_s and g are combined. This implies that μ_s' spectra alone might be insufficient to estimate the fat globule size and/or fat content and that accurate separation of μ_s and g would be required. This is, however, only feasible if the unscattered transmittance can be accurately measured, or if accurate diffuse reflectance and/or diffuse transmittance signals can be collected at very short source-detector distances (Watté et al., 2012; Kanick et al., 2012; Prahl, 2010; Sharma and Banerjee, 2003; Kienle et al., 2001).

As the milk fat globule quantity and size mainly determine the bulk scattering properties, they also have their impact on the measured signals. The diffuse reflectance, total transmittance and unscattered transmittance spectra for the 6 samples considered in Figure 2 are shown in Figure 3. As M_R and M_T are the integrated signals over all exit positions and exit angles of the light at respectively the reflectance and transmittance side of the sample, the similarity relation is valid and scattering can be described very accurately with only μ_s' (Prahl, 2010; Tuchin, 2007). Accordingly, the 2 groups that could be observed from the μ_s' spectra [Figure 2(c)] also appear in the M_R and M_T spectra [Figure 3(a) and (b)]. Consequently, the overall levels of the M_R and M_T spectra do not correlate well with the fat content of the samples. As μ_s is dominant over μ_a in the Vis/NIR range for raw milk, M_U is primarily influenced by μ_s. This can be clearly

observed in Figure 3(c). Moreover, as the M_U spectra are presented on a logarithmic scale, the plotted M_U spectra are very close to the inverse of the μ_s spectra [Figure 2(a)]. As a result, the 3 fat-content groups also clearly appear in the M_U spectra.

Relation between bulk optical properties and composition of raw milk

The fat globules and casein micelles are both important absorbing and scattering components in milk. Consequently, the correlation (R) between the content of milk fat and casein of all 60 samples [Table 1], and the measured μ_a and μ_s values have been calculated at each of the considered wavelengths [Figure 4]. High positive correlation coefficients of 0.751 and 0.632 – 0.762 were found between the fat content (% *v*/*v*) and the μ_a at 1220 and 1740 – 1770 nm. These are most likely related to respectively the second and first overtone stretch-vibrations of the CH_2-bonds (Šašić and Ozaki, 2000). Moreover, a negative correlation (R = -0.616) was found between the fat content and the water absorption at 1450 nm. As fat is an important part of the dry matter in milk, a negative correlation with the water content is obvious. The positive correlation ($R \geq 0.4$) between the fat content and μ_a from 700 until 1100 nm cannot be attributed to the absorption by fat. Moreover, it might be caused by the small cross-talk between μ_s and μ_a [Figure 1(a) Detail], as the correlation between the fat content and μ_s in that range is relatively high ($R \geq 0.839$) [Figure 4(b)].

Positive correlations of 0.378, 0.157 – 0.202 and 0.228 – 0.452 were found between the casein content (% *v*/*v*) and μ_a at 1250, 1580 – 1620 and 1670 – 1860 nm. These absorption peaks are probably related to respectively the first overtone of amide A + amide II vibrations, the overlapping first overtone of amide A and amide B vibrations and the first overtone stretch-vibrations of the CH-bonds in the protein side chains (Czarnik-Matusewicz et al., 1999). Similar to fat, casein followed a negative correlation (R = -0.252) with the water absorption at 1450 nm and an overall positive correlation with μ_a from 700 until 1100 nm. The latter might also be

explained as cross-talk from μ_s to μ_a, as in this wavelength range, μ_s has a considerable positive correlation ($R \geq 0.389$) with the casein content [Figure 4(b)]. Around 650 nm, the correlation between μ_a and both the fat and casein content drops, because the variability in μ_a at these wavelengths is mainly caused by a varying preservative concentration. Compared to milk fat, the casein content has an overall weaker correlation with μ_a, which can be explained by the smaller variability of crude protein in the analyzed milk samples [Table 1]. As the absorption peaks of fat, protein, water and/or other milk components overlap, it is not possible to get a perfect correlation between the absorption at a single wavelength and the concentration of a milk component. Combination of the absorption information present at different wavelengths through the use of multivariate calibration techniques could help to overcome this 'selectivity problem'. Since nearly no scattering effects are present in the μ_a spectra, accurate prediction models could potentially be built on these data without the need for empirical scatter corrections. Moreover, changes in the scattering properties would have (nearly) no impact on μ_a such that the predictions are expected to be robust.

In Figure 4(b) the correlation between the μ_s spectra and both the fat and casein content in raw milk is shown. An overall high positive correlation ($R \geq 0.766$) with the fat content (% *v*/*v*) was found, with the highest correlation ($R \geq 0.975$) in the 1300 to 1400 nm wavelength range. The very high correlation in the 1300 to 1400 nm region is probably because the size of the fat globules in raw milk has the least impact on μ_s at those wavelengths [Figure 2(a)]. Accordingly, μ_s will be more dependent on the fat content itself. For radiation wavelengths outside the 1300 – 1400 nm range, the size of the fat globules clearly affects μ_s, resulting in a lower correlation with the fat content itself.

A much lower correlation ($R = 0.276 – 0.556$) was found between the casein content (% *v*/*v*) and μ_s [Figure 4(b)]. This is probably because casein micelles contribute less to the light scattering in the Vis/NIR range (Aernouts et al., 2015; Bogomolov et al., 2012). The correlation,

however, increased with decreasing wavelength. Because casein micelles are small (10 – 500 nm) compared to the radiation wavelengths (550 – 1900 nm), the μ_s spectrum of casein micelles increases exponentially with decreasing wavelength towards the ultraviolet (UV) (Aernouts et al., 2015). Moreover, as the native fat globules in raw milk have a size (0.1 – 10 µm) similar to the radiation wavelengths, scattering increases with decreasing wavelength towards a maximum in the Vis/NIR, followed by a decrease towards the UV [Figure 2(a)]. As a result, the contribution of casein to the μ_s spectrum of raw milk increases for decreasing wavelengths in the UV/Vis, which confirms the correlations in Figure 4(b). These observations are supported by the findings from other Vis/NIR scattering experiments on raw milk (Kucheryavskiy et al., 2014; Bogomolov et al., 2013; Bogomolov and Melenteva, 2013; Bogomolov et al., 2012; Dahm, 2013). Moreover, it was found that scattering of raw milk at Vis wavelengths near the UV are more related to the casein content, while a better relation with the fat content was obtained towards the NIR.

In Figure 5, the relation between the fat content and μ_s at 3 radiation wavelengths (600, 1300 and 1700 nm) is provided. The solid lines in represent the linear fit between the fat content for all 60 raw milk samples and the respective μ_s at each of the 3 considered wavelengths. A poor correlation was found at 600 nm, while it was superior around 1300 nm. This was already clearly indicated in Figure 4(b). The fitted linear lines (solid) generally overestimate the μ_s at low (< 4% *v*/*v*) and high fat concentrations (> 8% *v*/*v*). This indicates that the relation between μ_s and the fat content is probably not linear, resulting from the effect of dependent scattering (Aernouts et al., 2014; Gaygadzhiev et al., 2008; Alexander et al., 2002). An earlier study on milk (Aernouts et al., 2015) showed that there was no significant effect of dependent scattering on the Vis/NIR μ_s spectra of raw milk if the fat content was below 3 – 4% (*v*/*v*). For these raw milk samples, the individual scattering processes will be independent and μ_s is expected to follow a linear increase with increasing fat content if the fat globule size is constant.

Nevertheless, if the fat content is above 4%, the scattering fat globules are close enough to influence the scattering by a neighboring fat globule. This generally results in a reduction of the bulk scattering coefficient spectra relative to those expected from the linear independent scattering relations (Aernouts et al., 2014; Gaygadzhiev et al., 2008; Alexander et al., 2002; Aernouts et al., 2015). To illustrate the effect of dependent scattering, a second linear curve (dashed line) was fitted between the μ_s and the fat content for the raw milk samples with 4% (*v*/*v*) fat or less [Figure 5]. Compared to the solid line (all data), the dashed line (independent scattering) resulted in a consistently higher slope. Moreover, the linear independent scattering fit (dashed line) generally overestimates the μ_s for fat contents above 4 – 5% (*v*/*v*), while this effect increases with increasing fat content (Aernouts et al., 2014; Gaygadzhiev et al., 2008; Alexander et al., 2002; Aernouts et al., 2015). At 600 nm wavelength, the difference between the slopes of the two linear fits is the largest [Figure 5 (a)]. This is probably related to the increased variability in μ_s for fixed fat contents [Figure 2(a)], additional to the effect of dependent scattering. Moreover, the increased variation is likely due to the effect of a varying fat globule size on μ_s, which is maximal for radiation wavelengths below 1100 nm [Figure 2(a)]. Accordingly, the effect of dependent scattering on the difference between slopes is inferior at these wavelengths.

As the effect of dependent scattering on μ_s is clearly present in the data [Figure 5], a non-linear model, taking into account this effect, would likely result in an improved fit with the data of μ_s versus the fat content (Aernouts et al., 2014; Gaygadzhiev et al., 2008; Alexander et al., 2002). Consequently, measurement of the μ_s at a single wavelength around 1300 nm could result in very accurate prediction of the fat content in raw milk samples.

The offset of the linear independent scattering fit (dashed line) gives the estimated average μ_s spectrum of these samples if no fat globules would be present [Figure 6]. Consequently, it relates to the average μ_s spectrum of the casein fraction in the raw milk

samples. In Figure 6, this offset is illustrated in function of the wavelength. Additionally, also the μ_s spectrum simulated for the casein fraction in a bulk milk sample (fat and crude protein content of respectively 4.52 and 2.65% *v*/*v*), as obtained from a previous study (Aernouts et al., 2015), was plotted. Although the offset-spectrum is the result of a fitting procedure on many diverse samples with a variable casein content [Table 1], there is a fairly good agreement between both curves.

CONCLUSION

The visible (Vis) and near-infrared (NIR) bulk optical properties of a set of 60 raw milk samples representative for milk from Flemish Holstein-Friesian cows have been measured on a double integrating spheres and unscattered transmittance setup. The variation in the absorption coefficient spectra was found to be clearly related to the composition of the milk samples, with clear influences of the water, fat and casein content. The bulk scattering coefficient spectra were found to be primarily influenced by the quantity and the size of the fat globules. A higher fat content results in an overall increase, while smaller fat globules produce steeper Vis/NIR bulk scattering coefficient spectra. Accordingly, the observed variation in the Vis/NIR bulk scattering coefficients was large. In the 1300 – 1400 nm wavelength range, the effect of the fat globule size on the bulk scattering coefficient of raw milk was found to be minimal, resulting in a strong positive correlation ($R \geq 0.975$) with the fat content. Moreover, the contribution of the fat content to the bulk scattering coefficient reduced towards the ultraviolet (UV), while the impact of the casein content increased. This could indicate the potential of UV scattering measurements for estimation of the casein content in raw milk. The anisotropy factor, on the other hand, is mainly influenced by the size of the fat globules and is nearly independent of the particle quantity. Moreover, larger milk fat globules cause more forward scattering of NIR light, which is represented by a higher anisotropy factor. As the fat and casein content had no

noticeable impact on the anisotropy factor, the variation in the anisotropy factor spectra of raw milk samples was rather limited.

The obtained information on the BOP of milk can be consulted in simulation studies to improve the insight in Vis/NIR light propagation in milk and other types of emulsions, which is essential for the optimal design of a Vis/NIR spectroscopic sensor that can accurately monitor the quality of raw milk. Moreover, this study indicates that, for the extraction of fat globule size from the scattering properties, a good separation between the bulk scattering coefficients and the anisotropy factors is essential. This can only be achieved with accurate unscattered transmittance measurements or multiple diffuse reflectance and/or diffuse transmittance measurements close to the point of illumination. Unscattered transmittance measurements of undiluted raw milk is, however, very challenging as very thin path lengths (< 200 µm) are required, the detector should be installed far (> 1 m) behind the sample and the unscattered transmittance signals are relatively weak. On the other hand, a small source-detector distance in diffuse reflectance and/or diffuse transmittance measurements results in a reduced penetration depth. Accordingly the minimal distance is limited as the sampled volume should be representative for the entire sample.

ACKNOWLEDGEMENTS

Ben Aernouts was funded as Ph. D. fellow of the Research Foundation-Flanders (FWO, grant 11A4813N). Rodrigo Watté, Robbe Van Beers and Tjebbe Huybrechts are funded by the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Flanders, respectively grants 101552, 131777 and 121611). The authors gratefully acknowledge IWT-Flanders for the financial support through the GlucoSens project (SB-090053).

REFERENCES

Aernouts, B., R. Van Beers, R. Watté, T. Huybrechts, J. Jordens, D. Vermeulen, T. Van Gerven, J. Lammertyn, and W. Saeys. 2015. Effect of ultrasonic homogenization on the Vis/NIR bulk optical properties of milk: measurements and simulations. *Colloids Surf. B Biointerfaces*. 126:510–519.

Aernouts, B., R. Van Beers, R. Watté, J. Lammertyn, and W. Saeys. 2014. Dependent scattering in intralipid phantoms in the 600-1850 nm range. *Opt. Express*. 22:6086–6098.

Aernouts, B., E. Polshin, J. Lammertyn, and W. Saeys. 2011. Visible and near-infrared spectroscopic analysis of raw milk for cow health monitoring: reflectance or transmittance? *J. Dairy Sci.* 94:5315–5329.

Aernouts, B., E. Zamora-Rojas, R. Van Beers, R. Watté, L. Wang, M. Tsuta, J. Lammertyn, and W. Saeys. 2013. Supercontinuum laser based optical characterization of turbid media in the 500-2250 nm range. *Opt. Express*. 21:32450–32467.

Alexander, M., L.F. Rojas-Ochoa, M. Leser, and P. Schurtenberger. 2002. Structure, dynamics, and optical properties of concentrated milk suspensions: an analogy to hard-sphere liquids. *J. Colloid Interface Sci.* 253:35–46.

Bogomolov, A., S. Dietrich, B. Boldrini, and R.W. Kessler. 2012. Quantitative determination of fat and total protein in milk based on visible light scatter. *Food Chem.* 134:412–418.

Bogomolov, A., and A. Melenteva. 2013. Scatter-based quantitative spectroscopic analysis of milk fat and total protein in the region 400–1100nm in the presence of fat globule size variability. *Chemom. Intell. Lab. Syst.* 126:129–139.

Bogomolov, A., A. Melenteva, and D. Dahm. 2013. Technical note: Fat globule size effect on visible and shortwave near infrared spectra of milk. *J. Near Infrared Spectrosc.* 21:435–440.

Cabassi, G., M. Profaizer, L. Marinoni, N. Rizzi, and T. Cattaneo. 2013. Estimation of fat globule size distribution in milk using an inverse light scattering model in the near infrared region. *J. Near Infrared Spectrosc.* 21:359–373.

Cattaneo, T., G. Cabassi, M. Profaizer, and R. Giangiacomo. 2009. Contribution of light scattering to near infrared absorption in milk. *J. Near Infrared Spectrosc.* 17:337–343.

Cen, H., R. Lu, and K. Dolan. 2010. Optimization of inverse algorithm for estimating the optical properties of biological materials using spatially-resolved diffuse reflectance. *Inverse Probl. Sci. Eng.* 18:853–872.

Czarnik-Matusewicz, B., K. Murayama, R. Tsenkova, and Y. Ozaki. 1999. of near-infrared spectra of complicated biological fluids by two-dimensional correlation spectroscopy: protein and fat concentration-dependent spectral changes of milk. *Appl. Spectrosc.* 53:1582–1594.

558 Dahm, D. 2013. Review: Explaining some light scattering properties of milk using
559 representative layer theory. *J. Near Infrared Spectrosc.* 21:323–339.

560 Erwin, R.E., and H.E. Randolph. 1975. Influence of mastitis on properties of milk. XI.
561 Fat globule membrane. *J. Dairy Sci.* 58:9–12.

562 Forsbäck, L., H. Lindmark-Månsson, A. Andrén, M. Akerstedt, L. Andrée, and K.
563 Svennersten-Sjaunja. 2010. Day-to-day variation in milk yield and milk composition at the
564 udder-quarter level. *J. Dairy Sci.* 93:3569–3577.

565 Forsbäck, L., H. Lindmark-Månsson, A. Andrén, M. Akerstedt, and K. Svennersten-
566 Sjaunja. 2009. Udder quarter milk composition at different levels of somatic cell count in cow
567 composite milk. *Animal*. 3:710–717.

568 Friggens, N.C., C. Ridder, and P. Løvendahl. 2007. On the use of milk composition
569 measures to predict the energy balance of dairy cows. *J. Dairy Sci.* 90:5453–5467.

570 Frisvad, J.R., N.J. Christensen, and H.W. Jensen. 2007. Computing the scattering
571 properties of participating media using Lorenz-Mie theory. *ACM Trans. Graph.* 26:60.

572 Gamm, U.A., S.C. Kanick, H.J.C.M. Sterenborg, D.J. Robinson, and A. Amelink.
573 2011. Measurement of tissue scattering properties using multi-diameter single fiber
574 reflectance spectroscopy: in silico sensitivity analysis. *Biomed. Opt. Express*. 2:3150–3166.

575 Gaygadzhiev, Z., M. Corredig, and M. Alexander. 2008. Diffusing wave spectroscopy
576 study of the colloidal interactions occurring between casein micelles and emulsion droplets:
577 comparison to hard-sphere behavior. *Langmuir*. 24:3794–3800.

578 Hamann, J., and V. Krömker. 1997. Potential of specific milk composition variables
579 for cow health management. *Livest. Prod. Sci.* 48:201–208.

580 ISO (International Organization for Standardization). 2000. Whole milk—
581 Determination of milk fat, protein and lactose content—Guidance on the operation of mid-
582 infrared instruments. International Standard ISO 9622:2000/IDF 141C:2000. International
583 Dairy Federation, Brussels, Belgium.

584 Kanick, S.C., V. Krishnaswamy, U. a Gamm, H.J.C.M. Sterenborg, D.J. Robinson, a
585 Amelink, and B.W. Pogue. 2012. Scattering phase function spectrum makes reflectance
586 spectrum measured from Intralipid phantoms and tissue sensitive to the device detection
587 geometry. *Biomed. Opt. Express*. 3:1086–100.

588 Kaniyamattam, K., and A. De Vries. 2014. Agreement between milk fat, protein, and
589 lactose observations collected from the Dairy Herd Improvement Association (DHIA) and a
590 real-time milk analyzer. *J. Dairy Sci.* 97:2896–2908.

591 Katz, G., Z. Schmilovitz, E. Maltz, M.I. Kutscher, M. Sarig, I. Halachmi, A. Hoffman,
592 H. Egozi, and E. Unar, inventors. 2003. Spectroscopic fluid analyzer. Kibutz Afikim (IL),
593 assignee. US Pat. No. 2003/0098969A1.

Katz, G., O. Shapira, L. Lemberskiy-Kuzin, and N. Pinsky, inventors. 2011. System and method for on-line analysis and sorting of milk coagulation properties. S.A.E. Afikim Milking System Agricultural Cooperative L.td, Kibutz Afikim (IL), assignee. US Pat. No. 8,072,596 B2.

Khankin, D., S. Mordechai, and S. Mark. 2012. Optimization Efficiency of Monte Carlo Simulation Tool for Evanescent Wave Spectroscopy Fiber-Optic Probe. *Adv. Opt. Technol.* 2012:1–5.

Kienle, A., F.K. Forster, and R. Hibst. 2001. Influence of the phase function on determination of the optical properties of biological tissue by spatially resolved reflectance. *Opt. Lett.* 26:1571–1573.

Kucheryavskiy, S., A. Melenteva, and A. Bogomolov. 2014. Determination of fat and total protein content in milk using conventional digital imaging. *Talanta*. 121:144–152.

Liu, Q., and N. Ramanujam. 2006. Sequential estimation of optical properties of a two-layered epithelial tissue model from depth-resolved ultraviolet-visible diffuse reflectance spectra. *Appl. Opt.* 45:4776–4790.

Logan, A., M. Auldist, J. Greenwood, and L. Day. 2014. Natural variation of bovine milk fat globule size within a herd. *J. Dairy Sci.* 97:4072–4082.

Løvendahl, P., C. Ridder, and N.C. Friggens. 2010. Limits to prediction of energy balance from milk composition measures at individual cow level. *J. Dairy Sci.* 93:1998–2006.

Luo, Y., H. Cui, X. Gu, R. Liu, and K. Xu. 2005. Determination of optimal source-detector separation in measuring chromophores in layered tissue with diffuse reflectance. *Chinese Opt. Lett.* 3:659–661.

McDowell, A. K. R. 1968. Fat testing of composite Milk samples with the Milko-tester. *J. Dairy Res.* 35:181–189.

Melfsen, A., E. Hartung, and A. Haeussermann. 2012. Potential of individual cow scatter correction for an improved accuracy of near infrared milk composition analysis. *J. Near Infrared Spectrosc.* 20:477–482.

Melfsen, A., E. Hartung, and A. Haeussermann. 2013. Robustness of near-infrared calibration models for the prediction of milk constituents during the milking process. *J. Dairy R.* 80:103–112.

Michalski, M.-C., B. Camier, V. Briard, N. Leconte, J.-Y. Gassi, H. Goudédranche, F. Michel, and J. Fauquant. 2004. The size of native milk fat globules affects physico-chemical and functional properties of Emmental cheese. *Lait*. 84:343–358.

Michalski, M.-C., J.-Y. Gassi, M.-H. Famelart, N. Leconte, B. Camier, F. Michel, and V. Briard-Bion. 2003. The size of native milk fat globules affects physico-chemical and sensory properties of Camembert cheese. *Lait*. 83:131–143.

630 Mizuno, K., M. Hatsuno, K. Aikawa, H. Takeichi, T. Himi, A. Kaneko, K. Kodaira, H.
631 Takahashi, and K. Itabashi. 2012. Mastitis is associated with IL-6 levels and milk fat globule
632 size in breast milk. *J. Hum. Lact.* 28:529–534.

633 Mulligan, F.J., L. O'Grady, D. a Rice, and M.L. Doherty. 2006. A herd health
634 approach to dairy cow nutrition and production diseases of the transition cow. *Anim. Reprod.*
635 *Sci.* 96:331–353.

636 Nielsen, N.I., T. Larsen, M. Bjerring, and K.L. Ingvartsen. 2005. Quarter health,
637 milking interval, and sampling time during milking affect the concentration of milk
638 constituents. *J. Dairy Sci.* 88:3186–3200.

639 Palmer, G., and N. Ramanujam. 2007. Use of genetic algorithms to optimize fiber
640 optic probe design for the extraction of tissue optical properties. *IEEE Trans. Biomed. Eng.*
641 54:1533–1535.

642 Pinsky, N., G. Katz, B. Sabbah, M.I. Kutscher, M. Sarig, Z. Merchav, and A. Gilboa,
643 inventors. 2013. System and method for analyzing fluids. Afikim Agricultural Cooperative
644 L.td., Kibutz Afikim (IL), assignee. US Pat. No. 8,446,582 B2.

645 Prahl, S.A. 2010. Everything I think you should know about inverse adding-doubling.
646 Accessed Dec. 20, 2014. http://omlc.ogi.edu/software/iad/iad-3-9-10.zip.

647 Šašić, S., and Y. Ozaki. 2000. Band assignment of near-infrared spectra of milk by use
648 of partial least-squares regression. *Appl. Spectrosc.* 54:1327–1338.

649 Schenkel, P., R. Samudrala, and J. Hinrichs. 2013. Thermo-physical properties of
650 semi-hard cheese made with different fat fractions: Influence of melting point and fat globule
651 size. *Int. Dairy J.* 30:79–87.

652 Sharma, D., A. Agrawal, L.S. Matchette, and T.J. Pfefer. 2006. Evaluation of a
653 fiberoptic-based system for measurement of optical properties in highly attenuating turbid
654 media. *Biomed. Eng. Online*. 5:49.

655 Sharma, S., and S. Banerjee. 2003. Role of approximate phase functions in Monte
656 Carlo simulation of light propagation in tissues. *J. Opt. A Pure Appl. Opt.* 5:294–302.

657 Steponavicius, R., and S. Thennadil. 2009. Extraction of chemical information of
658 suspensions using radiative transfer theory to remove multiple scattering effects: application
659 to a model two-component system. *Anal. Chem.* 81:7713–7723.

660 Steponavičius, R., and S.N. Thennadil. 2011. Extraction of chemical information of
661 suspensions using radiative transfer theory to remove multiple scattering effects: application
662 to a model multicomponent system. *Anal. Chem.* 83:1931–1937.

663 Steponavičius, R., and S.N. Thennadil. 2013. Full Correction of Scattering Effects by
664 Using the Radiative Transfer Theory for Improved Quantitative Analysis of Absorbing
665 Species in Suspensions. *Appl. Spectrosc.* 67:526–535.

Tuchin, V. V. 2007. Tissue Optics: Light Scattering Methods and Instruments for Medical Diagnosis. 2nd ed. SPIE Press, Washington, USA.

Vangroenweghe, F., H. Dosogne, and C. Burvenich. 2002. Composition and milk cell characteristics in quarter milk fractions of dairy cows with low cell count. *Vet. J.* 164:254–260.

Walstra, P., and R. Jenness. 1984. Dairy Chemistry and Physics. John Wiley and Sons, New York, USA.

Walstra, P., J. Wouters, and T. Geurts. 2006. Dairy Science and Technology. 2nd ed. Taylor & Francis Group, Boca Raton, USA.

Watté, R., B. Aernouts, and W. Saeys. 2012. A multilayer Monte Carlo method with free phase function choice. *Proc. SPIE*. 8429:84290S.

Zamora-Rojas, E., A. Garrido-Varo, B. Aernouts, D. Pérez-Marín, W. Saeys, Y. Yamada, and J.E. Guerrero-Ginel. 2014. Understanding Near-Infrared Radiation Propagation in Pig Skin Reflectance Measurements. *Innov. Food Sci. Emerg. Technol.* 22:137–146.

FIGURE TEXT

Figure 1: The mean, mean ± standard deviation (SD), minimum (Min) and maximum (Max) values for the bulk optical properties for 60 raw milk samples in the 550 – 1900 nm wavelength range: (a) bulk absorption coefficient μ_a; (b) bulk scattering coefficient μ_s; (c) anisotropy factor g; and (d) reduced scattering coefficient μ_s'.

Figure 2: The bulk scattering properties for 6 raw milk samples with a nearly constant crude protein content (2.36 – 2.49% *v*/*v*) and a varying fat content (% *v*/*v* in legend): (a) bulk scattering coefficient μ_s; (b) anisotropy factor g; and (c) the reduced scattering coefficient μ_s'.

Figure 3: The measured (a) diffuse reflectance M_R, (b) total transmittance M_T and (c) unscattered transmittance spectra M_U for 6 raw milk samples with a nearly constant crude protein content (2.36 – 2.49% *v*/*v*) and a varying fat content (% *v*/*v* in legend).

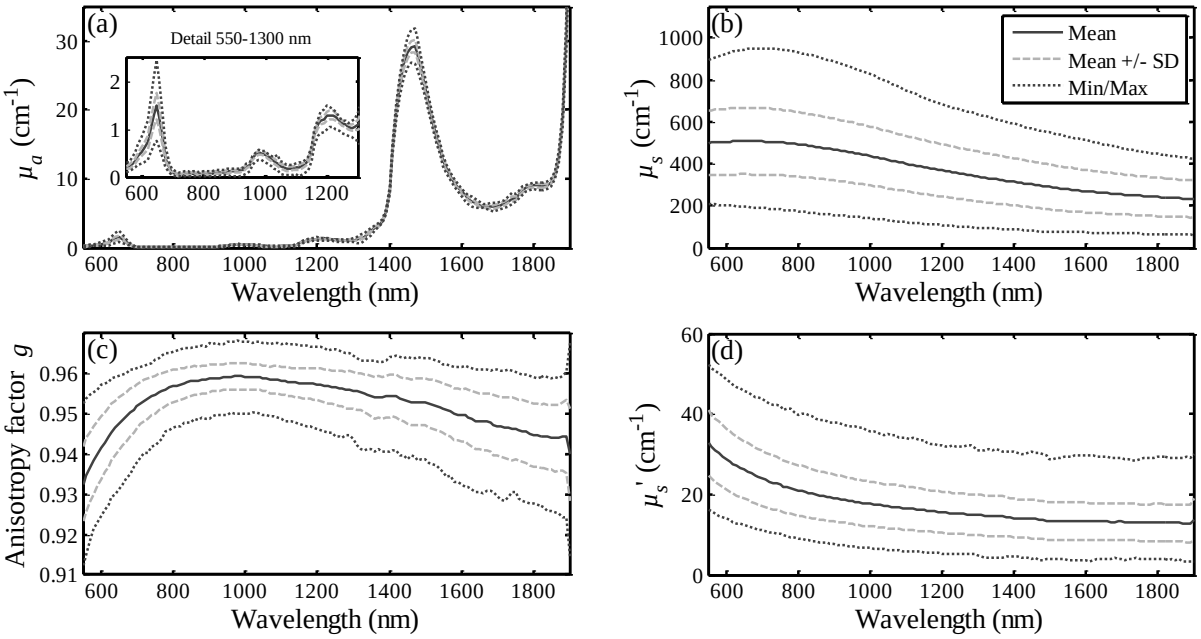
Figure 4: Correlation (R) between the content of milk fat and casein and (a) the bulk absorption coefficient μ_a and (b) the bulk scattering coefficient μ_s in function of the radiation wavelength.

Figure 5: Scatterplots of fat content (% *v*/*v*) versus the bulk scattering coefficient μ_s at (a) 600 nm, (b) 1300 nm and (c) 1700 nm. A linear curve is fitted to all the data points (solid lined) and to the data points for which the fat content was below or equal to 4% (*v*/*v*) (dashed line).

Figure 6: Simulated (*) μ_s spectrum for casein micelle fraction (2% *v*/*v*) in raw milk (Aernouts et al., 2015) and spectrum of offset values for the linear fit between the bulk scattering coefficient (μ_s) spectra and the fat content for samples with a fat content below 4% (*v*/*v*).

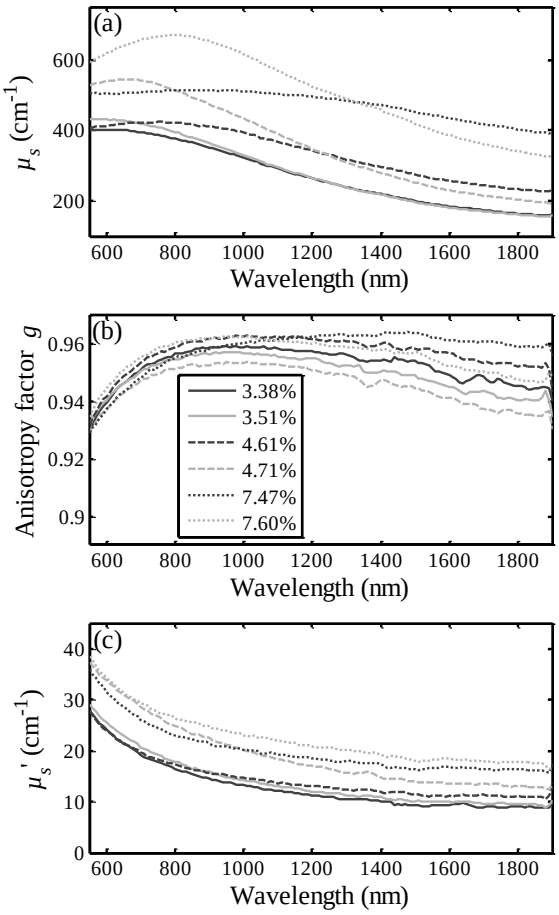
FIGURES

Figure 1



Aernouts: figure 1

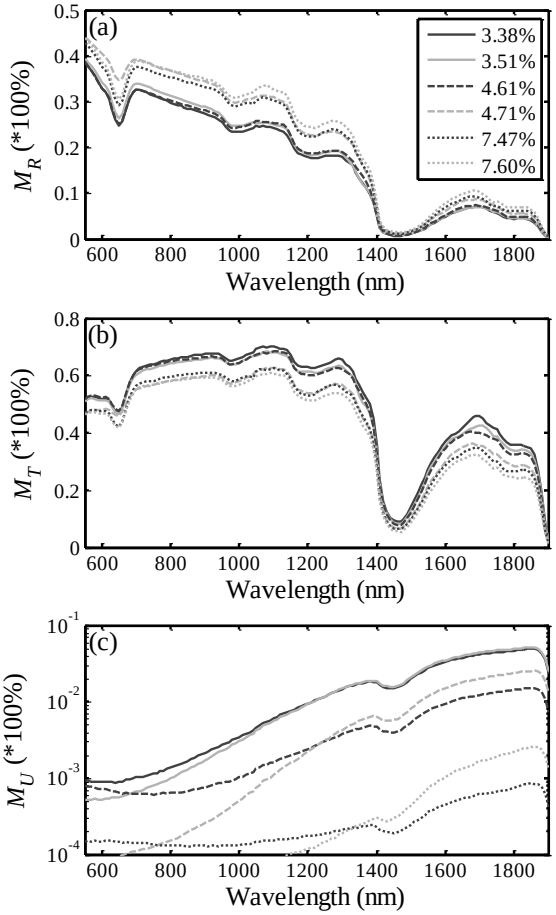
706 Figure 2



707

708 Aernouts: figure 2

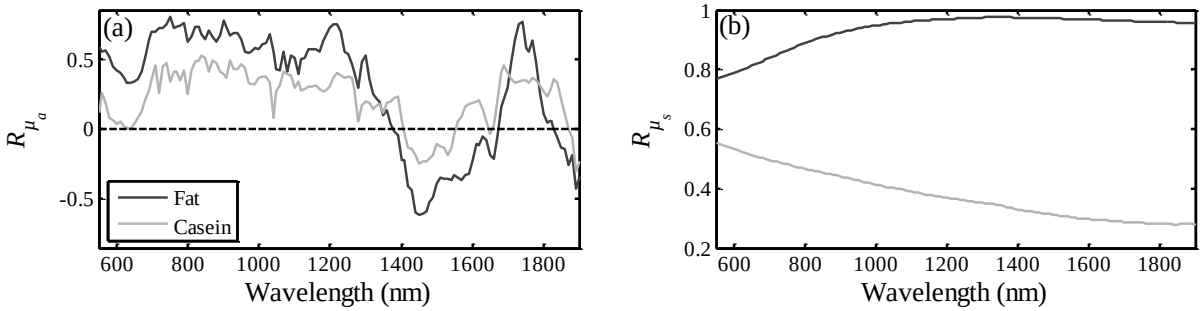
709 Figure 3



710

711 Aernouts: figure 3

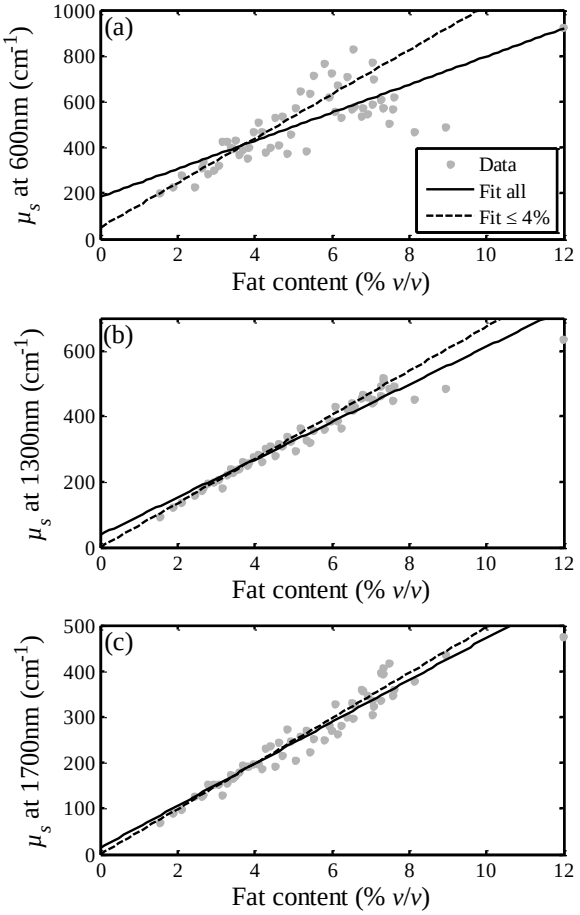
712 Figure 4



713

714 Aernouts: figure 4

715 Figure 5

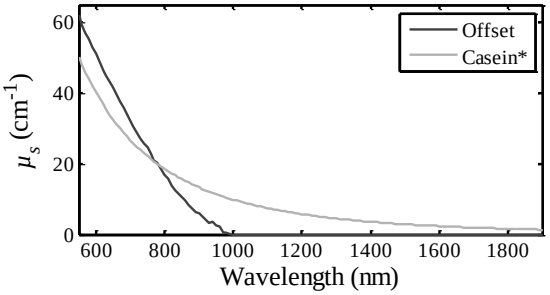


716

717 Aernouts: figure 5

718

719 Figure 6



720

721 Aernouts: figure 6

TABLES

Table 1 Basic statistics (all % *v*/*v*) on the main components in the 60 consulted raw milk samples.

Component	Mean	SD	Min	Max
Fat	5.24	2.01	1.52	12.0
Crude protein	2.72	0.424	2.09	3.73
Casein	2.05	0.424	1.58	2.82

739