

# Effect of fresh grass feeding, pasture grazing and organic/biodynamic farming on bovine milk triglyceride profile and implications for authentication

Edoardo Capuano · Rita Boerrigter-Eenling ·  
Anjo Elgersma · Saskia M. van Ruth

Received: 17 September 2013 / Revised: 22 November 2013 / Accepted: 30 November 2013 / Published online: 12 December 2013  
© Springer-Verlag Berlin Heidelberg 2013

**Abstract** In the present study, 113 tank milk samples were collected from 30 farms located in the Netherlands and analyzed for their triacylglycerol (TAG) and cholesterol profiles. Significant differences in the TAG profile between winter and spring–early summer milk were observed. The differences between milk from cows on pasture and stabled cows collected in spring–early summer were less remarkable than those between spring–early summer and winter milk. Classification models based on partial least square discriminant analysis of the TAG profile were developed for the prediction of fresh grass feeding, pasture grazing and organic farming. Milk from cows that had fresh grass in the daily ration could be distinguished from milk from cows that had no fresh grass with sensitivity and specificity values >85 %. However, the specificity reduced to 72 % when the samples collected in spring–early summer were considered only. Authentication of pasture grazing and of organic/biodynamic farming based on TAG profile proved difficult during the grazing period.

**Keywords** Milk · Authentication · Pasture · Grass · Organic · Triglycerides · Chemometrics · PLS-DA

## Introduction

Authentication of cattle feeding regimen and farming management system is gaining relevance due to the consumers'

increasing interest for practices that are environmental friendly and perceived as more respectful of animal welfare. In the dairy sector, pasture grazing and organic/biodynamic farming are known to confer distinct organoleptic properties and superior nutritional quality to milk [1] in the same time guaranteeing animal welfare and exploitation of local resources. Along with the organic or biodynamic certification, the “weidemelk” label has been recently introduced in the Netherlands for farm and processed milk (and other dairy products) produced from cows that have been allowed to graze on pasture at least 4 months per year, 6 h per day. Organic and “weidemelk” products are retailed at a higher price compared to conventional milk which might lead to mislabelling of dairy products for economic gain. Robust analytical methodologies that may underpin the paper-trail-based system of certification of products traceability are thus sought for.

Previous investigations proved that authentication of cows pasture feeding cannot be attained by the analysis of single biomarkers but rather by means of chemical or physical fingerprinting combined with chemometrics. For instance, Coppa et al. [2] showed that discrimination between milk from cows on pasture and stabled cows may be achieved by near infrared spectroscopy. However, the most suitable candidate for authentication of cows feeding regimen is the milk fat fraction since fatty acid (FA) and triacylglycerols (TAGs) profiles change very quickly following changes in the animals' diet. Different diets would indeed be characterized by a different FA profile and would differently modulate the microbial biohydrogenation activity in the rumen. TAGs account for approximately 98 % of milk fat thus constituting its major fraction. Milk fat TAGs vary considerably in molecular weight and degree of unsaturation [3]. This complexity is the direct result of the large number of FAs which are incorporated in the TAGs

E. Capuano (✉) · R. Boerrigter-Eenling · S. M. van Ruth  
RIKILT, Institute of Food Safety, Wageningen University  
and Research Centre, P.O. Box 230, 6700 AE Wageningen,  
The Netherlands  
e-mail: edoardo.capuano@wur.nl

A. Elgersma  
P.O. Box 323, 6700 AH Wageningen, The Netherlands

structure. Several studies have reported the seasonal effect on bovine milk TAG composition which stems from different feeding regimen implemented in winter and in spring-summer [4–6]. The effect of dietary lipids and specific feed supplementation on TAGs distribution has been also extensively addressed in the scientific literature [7–10]. However, the effect of fresh grass feeding and pasture grazing on bovine milk TAG composition has not. Analogously, information is scarce on the effect of organic/biodynamic farming on bovine milk TAGs composition.

The aim of this study was to explore the effect of fresh grass feeding, pasture grazing and organic farming on TAG profile of raw farm bovine milk and investigate the implications for the authentication of the cattle' management system. For that purpose, binary classifications models based on chemometric modeling have been developed for the authentication of fresh grass feeding, pasture grazing and organic farming based on cows' milk TAG profile.

## Materials and methods

### Study design

A total of 113 raw, tank milk samples were collected from 30 different farms located in the Gelderland region (6 farms) and in the Friesland/Groningen region (24 farms) in the Netherlands. Most of the farms provided 4 milk samples in different months, i.e., in April, May and June–July 2011 as well as in February 2012. The sampling scheme was setup so as to cover the spring feeding transition period (April and May) and the summer period (June–July) as well as the winter period (with cows on zero-grass ration). For practical reasons, it was not possible to plan an extra sampling round in autumn. The sampling was conducted in the framework of a larger project in which also other techniques/biomarkers were tested for the authentication of farming practice. The following information was collected at farms by means of interviews and questionnaires: (1) The hours spent by the cows on pasture at the moment

of sampling and (2) the composition of animals' diet at the moment of sampling. Data on the amount of fresh grass in the daily ration were also provided by the farmers. The amount of fresh grass of stabled cows was calculated from the acreage of cut land and the weight of fresh grass fed to the cows. The grass intake from grazing was estimated based on animal energy needs (based on milk production) minus what was fed next to the grassland. The milk samples were collected from the tanks by using 30-mL plastic screw-top containers after stirring the tank for at least 30 s. The milk samples were cooled (4 °C) immediately after sampling and frozen at –18°C within 6 h after sampling. Samples were categorized in five groups: Milk samples from cows that were at least 19-h outdoor on pasture on a daily basis at the time of sampling (group GP2); milk samples from cows were 6–9-h outdoor on pasture on a daily basis at the time of sampling (group GP1); milk samples from cows were indoors with fresh grass in the daily ration (group GI); milk samples collected in spring–early summer from cows that were indoors with no fresh grass in the diet (group NGI); milk samples collected in winter (group W). In group GP2, GP1 and GI, the estimated amount of fresh grass in the cows' daily ration varied from 36 to 94 %. All the milk samples of group W were from cows indoors all year-round with no fresh grass in the daily ration. Two certified organic farms and three certified biodynamic farms were included in the sample set. In Table 1, an overview of the study design is provided.

### TAG analysis

The TAG analysis was carried out according to International Standard ISO 17678/IDF 202 (milk fat—detection of foreign fats by gas chromatographic analysis of triglycerides). The TAGs were determined by gas chromatography by means of a Thermo Trace GC ultra model gas chromatograph (Thermo Scientific, Rodano, Italy), fitted with a flame-ionization detector (FID) and a column injector port. Chromatographic separation of TAGs was performed on a UltiMetal CP7532 (5 m × 0.53 mm × 0.17 μm, Varian,

**Table 1** Summary of the study design. Composition and characteristics of the different groups

Group	Hours on pasture per day	Min–max (average) % fresh grass per day	Samples collected per month <sup>a</sup>				Total <sup>a</sup>
			April	May	June	February	
GI	0	38–71 (54)	4	4	4	0	12
GP1	6–9	36–74 (51)	13 (4)	9 (1)	7	0	29 (5)
GP2	>19	52–94 (76)	4 (1)	9 (4)	11 (4)	0	24 (9)
NGI	0	0	7	6	6	0	19
W	0	0	0	0	0	29 (4)	29 (4)
Total <sup>a</sup>			28 (5)	28 (5)	28 (4)	29 (4)	113 (18)

<sup>a</sup> In parenthesis the number of organic/biodynamic farms

The Netherlands). TAGs were identified by their retention times compared to those found in the TAG standard mixture (certified reference materials BCR 632a and BCR632b, Institute for reference materials and measurements, IRRM, Geel, Belgium). All milk samples were analyzed in triplicate, and results were expressed as normalized peak areas (%). Data used for modeling were the average values of the three replicates of each milk sample.

#### Fatty acids analysis

Approximately 15 g of milk were mixed with 24 mL of a dichloromethane–ethanol solution (2:1 ratio, v/v) in a 40-mL glass test tube. The mixture was shaken with a vortex for 5 min and then centrifuged for 15 min (2,500 g at 48 °C). The upper organic phase was removed with a pipette, and further 15 mL of the dichloromethane–ethanol solution was added to the test tube and then the mixture vortexed and centrifuged for 15 min (2,500 g at 48 °C). The combined upper organic phases containing the milk fat was filtered, and dichloromethane–ethanol was evaporated until dryness under a stream of nitrogen.

The FAMES were obtained using a two-step methylation procedure with sodium methoxide (0.5 N) and boron trifluoride–methanol complex (20 %) as previously described [11]. The FAME analysis was carried out by means of a Focus GC gas chromatograph (Thermo Scientific, Italy), fitted with a flame-ionization detector (GC–FID) and split-splitless injector port, set at 280 and 250 °C, respectively, and a split ratio of 1:30. Chromatographic separation of FAMES was performed on a CP 7419 FAME capillary column (50 m × 0.25 mm inner diameter; Varian, Palo Alto, CA). Helium (1 mL min<sup>-1</sup> flow) was used as a carrier gas. The oven was programmed as follows: initial temperature, 100 °C, increased at 5 °C/min to 230 °C and held for 9 min. The sample volume injected was 1 µL. FAMES were identified by comparing their retention times with that of the available FAME standards (Supelco 37 Component FAME mix, Supelco, St. Louis, MO and standards for *cis*-9, *trans*-11, C18:2 and *trans*-11 C18:1 methyl esters, Sigma-Aldrich, St. Louis, MO). All milk samples were analyzed in triplicate, and results were expressed as normalized peak areas (%). Data used were the average values of the three replicates of each milk sample.

#### Statistical analysis

First, the normality of TAGs distribution within the groups GI, GP1, GP2, NGI and W was checked by means of a Shapiro–Wilks test for normality ( $p < 0.05$  was considered significant). Since TAGs distribution appeared to be non-normal in the groups, a Kruskal–Wallis test for group comparisons was performed among the groups NGI, GI, GP1,

GP2 and W ( $p < 0.05$  was considered significant). Correlation between pairs of TAGs was calculated by means of Spearman correlation test (2-tails,  $p < 0.05$  was considered significant). SPSS version 19.0 (IBM Statistics Inc., Armonk, NY) was used to perform all those tests.

For multivariate analysis, principal component analysis (PCA) of the TAGs data was performed with the 113 samples to explore the presence of outliers, the presence of any natural clustering in the data and to visualize the correlation between TAGs and FAs. Pirouette 4.5 (Infometrix, Seattle, USA) was used to develop the PCA models. Partial least squares discriminant analysis (PLS-DA) was performed to develop classification models to verify the presence of fresh grass in the animal diet, pasture grazing and organic/biodynamic management system. In a PLS-DA model, new variables (latent variables or factors) are created from the combination of the original variables in order to find the maximum correlation between the new variables and the class variable and, as a result, the maximum separation among the classes. The performance of the PLS-DA algorithm was assessed using a stratified random resampling approach including internal and external model validation as described in Tres et al. [12]. Briefly, the resampling approach consisted in four steps as follows: (i) A training set (75 % of the samples of each class) is randomly selected and the data processed, (ii) the optimal number of latent variables is computed using the routine of Boulesteix [13], (iii) the preprocessed dataset is cross-validated and (iv) the optimized model is validated with the test set (the remaining 25 % of the samples). These four steps were repeated 300 times. Thus, 300 data subsets were created, and each of them was subjected to the three data preprocessing techniques, namely mean-centering, auto-scaling (scaling to unit variance) and direct orthogonal signal correction (DOSC). In DOSC, the information that is not related to the response variable (or class membership) is largely ignored [14]. Then, each of the 900 models was internally and externally validated. The optimal number of latent variables to be included in each model was selected in order to give the least number of misclassifications among the samples of the training set. Sensitivity (i.e., % of samples of class 1 correctly assigned to the class 1) and the specificity (i.e., % of samples of class 2 correctly assigned to class 2) were calculated for the training and the validation set and then averaged on the whole procedure yielding mean sensitivity and specificity. A variable selection approach was used to identify which variable (TAG) most contributed to the PLS-DA classification models as described in Tres et al. [12]. Finally, we assessed the frequency of misclassification by determining the number of instances where each sample in the dataset was incorrectly classified during external validation and dividing that number by the

number of instances where that sample was present in the external validation set.

R version 2.12.2 ([www.R-project.org](http://www.R-project.org)) [15] was used to develop the PLS-DA classification models using built-in functions and the package “pls.genomics” for various PLS-DA algorithms. For DOSC preprocessing, the MATLAB code provided by the Biosystems Data Analysis Group of the University of Amsterdam (<http://www.bdagroup.nl/>) was rewritten to R code.

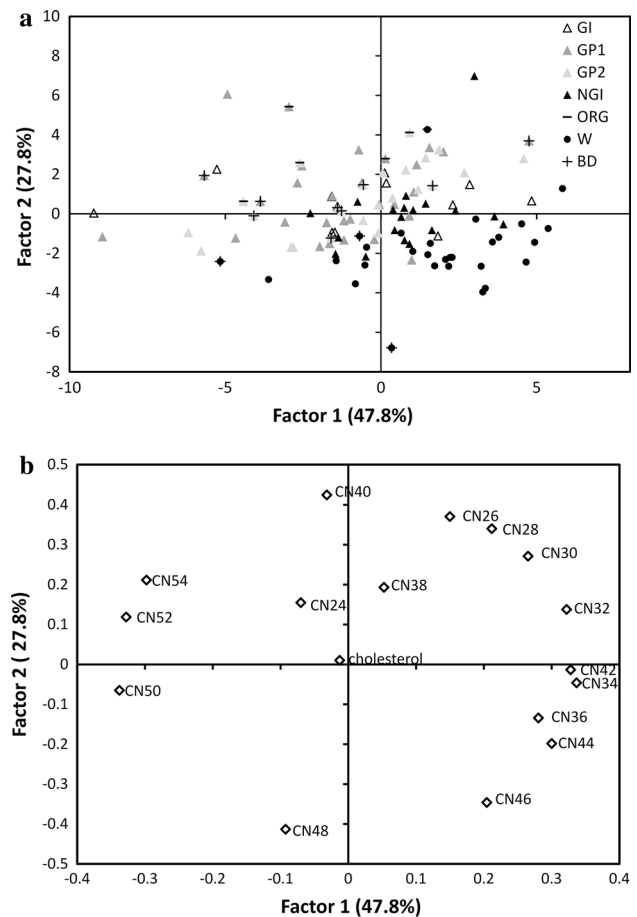
## Results and discussion

### TAG composition of milk

A total of 16 TAG groups plus cholesterol were quantified by GC–FID. Each TAG group is constituted by all the triglycerides with a certain total number of acyl carbon number (similar molecular weight), e.g., CN54 is the sum of all the TAG containing 54 acyl carbon atoms. The chromatographic method that was used cannot resolve a TAG group in its individual triglycerides components. The average TAG composition was in agreement with literature data [16]. The TAG profiles that were obtained in this study were bimodal, which is typical of cows' milk, with clear maxima at CN38 and CN50 (Table 1). The most abundant TAGs (expressed as mean  $\pm$  SD) were CN38 ( $13.16 \pm 0.38$  g 100 g fat<sup>-1</sup>), C36 ( $11.17 \pm 0.79$  g 100 g fat<sup>-1</sup>), CN40 ( $10.42 \pm 0.45$  g 100 g fat<sup>-1</sup>) and CN50 ( $10.45 \pm 0.95$  g 100 g fat<sup>-1</sup>). The average cholesterol content was  $0.37 \pm 0.03$  g 100 g fat<sup>-1</sup> which was higher than reported in European bulk milk by Precht [17] but in line with other literature data [18]. Significant positive correlations were found between TAGs within each of the groups CN26–30 ( $r > 0.804$ ,  $p = 0.000$ ), CN32–36 ( $r > 0.559$ ,  $p = 0.000$ ), CN42–46 ( $r > 0.609$ ,  $p = 0.000$ ) and CN50–54 ( $r > 0.731$ ,  $p = 0.000$ ) and significant negative correlation between any TAG from the groups CN50–54 and any TAG from groups CN32–36 ( $r > 0.560$ ,  $p = 0.000$ ), CN42–46 ( $r > 0.445$ ,  $p = 0.000$ ) as well as between CN50 and any TAG from groups CN26–30 ( $r > 0.438$ ,  $p = 0.000$ ). Cholesterol content was significantly correlated only with the content in CN24 ( $r = 0.451$ ,  $p = 0.000$ ) and CN38 ( $r = -0.311$ ,  $p = 0.01$ ). A relatively higher positive correlation between cholesterol content and short-chained triglycerides, especially CN26, CN28 and CN30, has been previously reported [17] which was not observed in the present study.

### Exploring the TAGs data: PCA and inferential statistics

First, a PCA model was created to check natural clustering in the data and the presence of outliers. The data matrix



**Fig. 1** First and second principal components of the PCA scores plot (panel a) and PCA loadings plot (panel b) of triacylglycerols profile + cholesterol (auto-scaled data) for farm milk samples (the amount of explained variance is provided in parenthesis). Group GI = milk from cows that have been fed grass indoors; GP1 = milk from cows that have been on pasture 6–9 h per day; GP2 = milk from cows that have been on pasture >19 h per day; NGI = milk from cows not fed fresh grass collected in spring–early summer; W = milk collected in winter; ORG = organic milk samples; BD = biodynamic milk samples

consisted of 113 rows (samples) and 17 variables (16 TAG groups and cholesterol). The first two principal components (PCs) explain approximately 76 % of the total variance. The PCA scores plot of the auto-scaled data is depicted in Fig. 1a. From the visual inspection of the PCA scores plot, the most clear difference was between milk samples collected in February and those collected in spring–early summer. Winter samples are mostly enclosed in the SE quadrant with positive scores along PC 1 and negative scores along PC 2. In Fig. 1b, the PCA loadings plot visually illustrates the correlations among TAGs and TAG groups discussed in “TAG composition of milk” section and the compositional differences among the different groups. The outcome of the PCA analysis is confirmed by the results of the Kruskal–Wallis test reported in Table 2. The distribution of 12 TAGs

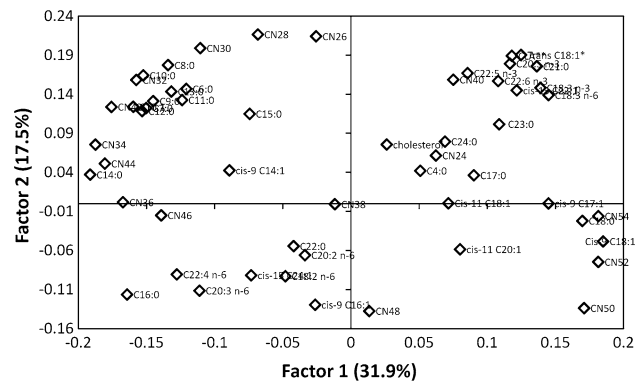
**Table 2** Triacylglycerols composition (g 100 g fat<sup>-1</sup>) in the group GI (milk from cows that have been fed grass indoors), GP1 (milk from cows that have not outdoors on pasture 6–9 h per day), GP2 (milk from cows that have not outdoors on pasture >19 h per day), NGI (milk from cows not fed fresh grass collected in spring) and W (samples collected in winter)

TG	GI	GP1	GP2	NGI	W
CN24	0.09 <sup>ab</sup>	0.09 <sup>a</sup>	0.10 <sup>a</sup>	0.10 <sup>a</sup>	0.08 <sup>b</sup>
Cholesterol	0.38 <sup>a</sup>	0.37 <sup>a</sup>	0.38 <sup>a</sup>	0.37 <sup>ab</sup>	0.35 <sup>b</sup>
CN26	0.31 <sup>ab</sup>	0.32 <sup>a</sup>	0.33 <sup>a</sup>	0.31 <sup>ab</sup>	0.28 <sup>b</sup>
CN28	0.70	0.72	0.73	0.72	0.66
CN30	1.28	1.28	1.31	1.34	1.25
CN32	2.60	2.58	2.63	2.72	2.71
CN34	5.98 <sup>ab</sup>	5.84 <sup>a</sup>	5.89 <sup>ac</sup>	6.34 <sup>bc</sup>	6.56 <sup>b</sup>
CN36	10.96 <sup>ab</sup>	10.70 <sup>b</sup>	10.66 <sup>b</sup>	11.49 <sup>ac</sup>	11.93 <sup>c</sup>
CN38	13.30 <sup>b</sup>	13.15 <sup>ab</sup>	12.93 <sup>a</sup>	13.28 <sup>b</sup>	13.20 <sup>ab</sup>
CN40	10.53 <sup>ab</sup>	10.61 <sup>ab</sup>	10.64 <sup>a</sup>	10.30 <sup>bd</sup>	10.07 <sup>cd</sup>
CN42	7.43 <sup>ab</sup>	7.27 <sup>a</sup>	7.48 <sup>ab</sup>	7.64 <sup>ab</sup>	7.86 <sup>b</sup>
CN44	7.23 <sup>a</sup>	7.06 <sup>a</sup>	7.26 <sup>a</sup>	7.50 <sup>ab</sup>	7.95 <sup>b</sup>
CN46	8.23 <sup>a</sup>	8.14 <sup>a</sup>	8.33 <sup>a</sup>	8.47 <sup>a</sup>	8.94 <sup>b</sup>
CN48	9.81	9.80	9.91	9.91	10.13
CN50	10.61 <sup>ab</sup>	10.82 <sup>a</sup>	10.69 <sup>a</sup>	10.32 <sup>ab</sup>	9.91 <sup>b</sup>
CN52	7.83 <sup>ab</sup>	8.26 <sup>a</sup>	7.87 <sup>ab</sup>	7.03 <sup>bd</sup>	6.29 <sup>cd</sup>
CN54	2.74 <sup>ab</sup>	2.98 <sup>a</sup>	2.88 <sup>a</sup>	2.18 <sup>bc</sup>	1.81 <sup>c</sup>

Different letters within a row indicate significant difference among triacylglycerols distribution within the 5 groups (Kruskal–Wallis test for groups comparisons,  $p \leq 0.05$ )

as well as of cholesterol is significantly different among the 5 groups. The distribution of 11 TAGs and cholesterol is significantly different between the samples collected in February and at least one of the groups of milk samples collected in spring–early summer.

Seasonal differences in TAG profile has been already reported in the scientific literature [4–6]. Differences in animals diet are responsible for the observed differences with feeding regime implements in spring–early summer richer in grass related feeding, e.g., fresh grass and grass silage compared to feeding regime implemented in winter. In winter, when fresh grass is not available, maize silage and concentrates make up most of the animals daily ration. The different content in CN54 was especially expected. The CN54 group is almost exclusively made up of triglycerides incorporating three C18 FA [19] which content is higher in milk from cows fed fresh grass because of  $\alpha$ -linolenic acid being the predominant FA in fresh grass [20, 21]. Similarly, TAG CN52 is also made up of triglycerides that contains two C18 FAs, and thus, its relative content is expected to be higher in spring–early summer milk. The medium molecular weight triglycerides CN34 and CN36 incorporate mostly short- and medium-chain saturated FAs (butyric, myristic, palmitic) and stearic acid which



**Fig. 2** First and second principal components of the PCA loadings plot of triacylglycerols + cholesterol and fatty acid profiles (auto-scaled data, the amount of explained variance is provided in parenthesis). \*Sum of all *trans*-C18:1 fatty acids; \*\**cis*-9, *trans*-11 conjugated linoleic acid

concentration is higher in winter [16]. It is thus not surprising that those TAG groups were higher in winter milk. Finally, TAG CN42–46 were also more abundant in winter milk compared to spring–early summer milk. According to Gresti et al. [19], those TAG are mainly made up of a saturated FA esterified in C1, a saturated FA (mainly palmitic acid) or oleic acid in C-2 and mostly oleic acid in C-3. To visually illustrate the correlation between the TAG and the FAs content in our samples, a PCA model was created with the combined TAG and FA data. The PCA loadings plot is reported in Fig. 2. The first two PCs explain almost 50 % of the total variance. Variables that are close in the PCA loadings plot, especially along the first component, are highly correlated. Figure 2 shows that TAG50–54 are highly positively correlated with oleic acid (*cis*-9 C18:1) and stearic acid (C18:0) and negatively correlated with short-chain saturated fatty acids; CN30–34 and CN42–44 are highly positively correlated with short-chain saturated FAs (C6:0–C12:0), whereas CN34, CN36 and CN44 are highly positively correlated with myristic acid (C14:0). Finally, CN40 is highly positively correlated with several long-chain polyunsaturated FAs and negatively correlated with palmitic acid (C16:0).

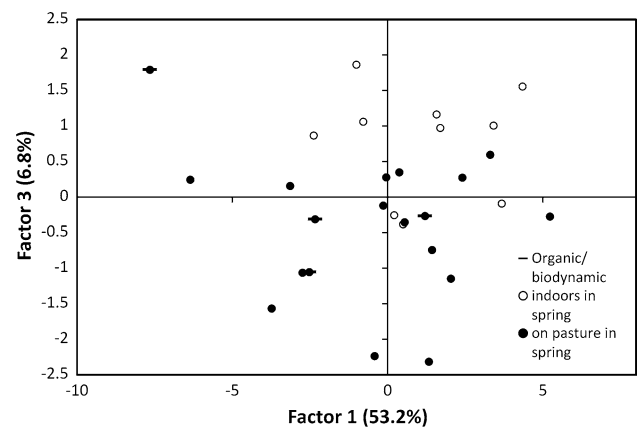
Interestingly, cholesterol content was higher in spring–early summer compared to winter milk, whereas no effect of fresh grass feeding/pasture was observed. A decrease in milk cholesterol under energetic underfeeding conditions and during the pasture period has been reported [17]. Conversely, a mean cholesterol content of 310 mg 100 g lipids<sup>-1</sup> in German summer milk compared to 240 mg 100 g lipids<sup>-1</sup> winter summer milk was also reported [22].

Since the presence of fresh grass and pasture grazing are among the causes for the differences in the TAG profile between spring–early summer and winter samples, we expected to find analogous differences between the milk



samples from cows with and without fresh grass in their daily ration collected in spring–early summer. From the PCA scores plot of Fig. 1, it appears that an extensive overlap occurs between group NGI and groups GP1 and GP2. From Table 2, it can be inferred that the distributions of CN36 and CN54 are significantly different in groups GP1 and GP2 compared to group NGI. The distributions of two TAGs, namely CN34 and CN52, were significantly different between groups GP1 and NGI, whereas the distribution of CN38 and CN40 was significantly different between groups GP2 and NGI. The higher content in CN52 and CN54 and the lower content in CN34, CN36 and CN38 in the milk of cows with fresh grass in the diet is consistent with a diet rich in fresh grass compared to milk from cows without fresh grass as discussed above. From Fig. 1a and Table 2, it is evident that the differences between milk from cows on pasture and spring–early summer milk from cows indoors without fresh grass in the daily ration are less remarkable than those between winter and spring–early summer milk. This would indicate that the seasonal differences observed in the TAG profile are not exclusively related to the presence/absence of fresh grass in the diet but relate to the animal diet in a more complex way, for example to the presence of grass silage in the spring–early summer diet and higher concentrates and maize silage in the winter diet. When milk from cows with fresh grass in the diet was considered, no difference appears in the PCA scores plot between group GP1 and GP2 which is also confirmed by the results reported in Table 2 (no TAG distribution was significantly different between the two groups from their pairwise comparison), nor within group GI and group GP1 or GP2. In the PCA scores plot of Fig. 1a, the organic/biodynamic milk samples are also indicated. In the plot, the organic/biodynamic samples are mostly mixed with the conventional samples and do not form a distinct cluster which suggest limited differences between the two management systems.

The sample set constituted by the samples collected in winter was also subjected to an independent multivariate analysis. A PCA model was created with the 29 samples collected in February. The PCA scores of the first two PCs plot did not show any clear clustering of the samples. The four samples from organic/biodynamic farms were quite distant from the main sample cloud, but they were scattered over the PCA scores plot and did not form a homogeneous group. However, a moderate discrimination between milk from farms with cows on pasture during spring–early summer and milk from farms implementing an indoors management system year-round can be observed along the third PC (Fig. 3). Cholesterol and CN24 showed the highest loadings along that component, whereas CN36 and CN38 showed the lowest. This difference would suggest that milk fat from grazing cows shows differences



**Fig. 3** First and third principal components of the PCA scores plot of triacylglycerols profile + cholesterol (auto-scaled data) for 29 farm milk samples collected in February (the amount of explained variance is provided in parenthesis)

not only in spring–early summer but year-round. Consistent differences in the feeding strategies between the two housing systems were observed, with lower ratio of grass silage + hay relative to concentrate + maize in those farms implementing zero-grazing systems year-round.

#### Binary classification models

One of the aims of the present study was to investigate the possibility to exploit the variability in TAG profile induced by diet for authentication purposes. In particular, three independent sets of binary classification models were created to predict whether a new unknown milk sample (1) was from cows that have been fed fresh grass or not, (2) was from cows that have been kept indoors or grazing on pasture and (3) was from conventional or organic/biodynamic farming.

In the first set of PLS-DA models, the samples of group NGI and W were pooled in the class NO GRASS, whereas the remaining three groups were pooled in the class GRASS. The three preprocessing techniques gave very similar results, with auto-scaling and DOSC giving sensitivity (% of corrected classification in group GRASS) and specificity (% of correct classification in class NO GRASS) values of 95 and 91 % in internal validation and 91 and 86 % in external validation which makes the performance of the PLS-DA model satisfactorily (Table 3). However, when the results are recalculated over the spring–early summer samples only, a specificity of 72 % was achieved, whereas the sensitivity was not affected since milk samples from grass-fed cows were obviously all collected in spring–early summer. This in fact limits the practical applicability of this approach for the verification of fresh grass feeding. The first five variables (TAGs) that contributed most to the discrimination between the classes in the PLS-DA models

**Table 3** Mean sensitivity and specificity of PLS-DA models during internal and external validation established from 300 realizations of the PLS-DA model using several methods of data preprocessing

Preprocessing	Mean sensitivity (%)	Mean specificity (%)
Model 1: Fresh grass versus no fresh grass		
Internal validation		
Mean-centering	93	90
Auto-scaling	95	91
DOSC	95	91
External validation		
Mean-centering	90	85
Auto-scaling	90	86
DOSC	91	85
Model 2: Pasture grazing versus indoors feeding		
Internal validation		
Mean-centering	90	87
Auto-scaling	90	85
DOSC	93	90
External validation		
Mean-centering	82	81
Auto-scaling	82	80
DOSC	84	81
Model 3: organic/biodynamic farming versus conventional farming		
Internal validation		
Mean-centering	96	79
Auto-scaling	96	82
DOSC	97	90
External validation		
Mean-centering	91	57
Auto-scaling	90	56
DOSC	89	60

were: CN36, CN54, CN52, CN40 and CN34. The results of this analysis are in agreement with the outcome of the PCA analysis described in “[Exploring the TAGs data: PCA and inferential statistics](#)” section and depicted in the PCA loadings plot in Fig. 1b.

The second set of PLS-DA classification models was created for the discrimination between milk samples according to whether or not cows have been grazing on pasture. Groups NGI, W and GI were grouped in the class INDOORS and groups GP1 and GP2 in the class PASTURE. The DOSC preprocessing of the data proved more satisfactorily in terms of sensitivity and specificity compared to mean-centering and auto-scaling (Table 3). Overall, the discrimination between milk from cows on pasture or indoors is less accurate than the discrimination between milk from cows with or without fresh grass in the diet. An average percentage of correct prediction of 93 and 90 % in internal validation and 84 and 81 % in external validation was achieved. Also in this case, when the prediction is

limited to the spring–early summer samples, the specificity values further reduce to approximately 63 % that make the approach of limited use in practice. The analysis of the misclassified samples showed that, within the class INDOORS, the frequency of misclassification was much higher for milk from cows fed fresh grass indoors (group GI) compared to the milk from cows without fresh grass (group NGI). This was expected because TAG profile of samples of group GI was on average more similar to that of milk from cows on pasture compared to stabled cows (Fig. 1a).

For the verification of the organic/biodynamic management system, since remarkable differences exist between the feeding regime implemented in spring–early summer and that implemented in winter, only milk samples collected in spring–early summer were included in the binary classification models. A set of binary classification models with samples collected in winter was not developed because of the limited number of samples collected in February. From Table 3, it appears that the DOSC preprocessing of the data gave the best prediction results in internal and external validation compared to auto-scaling and mean-centering. However, it is clear from the low specificity values in external validation (60 %) that the correct classification of the organic/biodynamic milk samples was difficult. We already reported the results of a PLS-DA model for the discrimination between organic and conventional bovine farm milks [23]. In that study, 85 and 89 % of the organic and conventional samples were correctly classified by the model in internal validation, which is comparable to the results obtained in the present study, but no external validation of the PLS-DA model was performed.

## Conclusions

In the present study, we provided insight into the effect of fresh grass feeding, pasture grazing and organic/biodynamic farming on bovine farm milk TAG profile. We demonstrated that winter farm milk showed a TAG profile that was remarkably different from that of spring–early summer milk because of the seasonal differences in feeding strategies. In spring–early summer, fresh grass feeding and pasture grazing contributed to the differences in the TAG profile between milk from stabled cows and milk from cows on pasture, increasing the relative concentrations of CN52, CN54 and CN40 in the latter. PLS-DA binary classification model appeared to be successful in the prediction of the presence of fresh grass in the animal’s diet, with sensitivity and specificity values  $\approx 90$  % when milk collected in spring–early summer, and winter was considered but the specificity values drastically reduced to 72 % when the prediction was limited only to samples collected in spring–early summer. The difference in the TAG profile of milk

from cows indoors and that from cows on pasture is related to the different share of fresh grass in the animals' diet. The authentication of pasture grazing based on TAG profile is therefore less accurate than the authentication of fresh grass feeding because of the possibility of feeding fresh grass to cows indoors. Analogously, authentication of organic farming seems to be limitedly feasible based on TAG profiles in spring–early summer or in winter. It is possible that more detailed information on triglyceride composition such as, for instance, milk fat fractionation into more specific triacylglycerol classes on the basis of molecular weight and degree of unsaturation or even the separation of individual triglycerides may provide more insight into the animal diet/farming systems.

**Acknowledgments** This study has been funded by the Dutch Ministry of Economic Affairs and Innovation. The authors gratefully acknowledge the farmers for providing milk samples and Adriana Sterian for having carried out part of the analysis. They also acknowledge Grishja van der Veer for the support on the statistical analysis of the data.

**Conflict of interest** None.

**Compliance with Ethics Requirements** This article does not contain any studies with human or animal subjects.

## References

- Elgersma A (2012) New developments in The Netherlands: dairies reward grazing because of public perception. *Grassland Sci Eur* 17:420–422
- Coppa M, Martin B, Agabriel C, Chassaing C, Sibra C, Constant I, Graulet B, Andueza D (2012) Authentication of cow feeding and geographic origin on milk using visible and near-infrared spectroscopy. *J Dairy Sci* 95:5544–5551
- Belitz HD, Grosch W, Schieberle P (2009) *Food chemistry*, 4th edn. Springer Verlag, Berlin
- Bomaz S, Novak G, Parmentier M (1992) Seasonal and regional variation in triglyceride composition of French butterfat. *J Am Oil Chem Soc* 69:1131–1135
- Precht D, Frede E (1996) Determination of the solid fat content in milk fat by gas chromatographic triglyceride analysis. *Fat Sci Technol* 96:324–330
- Hinrichs J, Heinemann U, Kessler HG (1992) Differences in the composition of triglycerides in summer and winter milk fat. *Milchwissenschaft* 47(8):492–495
- Banks W, Clapperton JL, Muir DD, Girdler AK (1989) Whipping properties of cream in relation to milk composition. *J Dairy Res* 56:97–105
- Murphy JJ, McNeill GP, Connolly JF, Gleeson PA (1990) Effect on cow performance and milk fat composition of including full fat soyabeans and rapeseeds in the concentrate mixture for lactating dairy cows. *J Dairy Res* 56:295–306
- DePeters EJ, German JB, Taylor SJ, Essex ST, Perez-Monti H (2001) Fatty acid and triglyceride composition of milk fat from lactating holstein cows in response to supplemental canola oil. *J Dairy Sci* 84:929–936
- Palmquist DL, Beaulieu AD, Barbano DM (1993) Feed and animal factors influencing milk fat composition. *J Dairy Sci* 76:1753–1771
- Tres A, Van Ruth SM (2011) Verification of organic feed identity by fatty acid fingerprinting. *J Agric Food Chem* 59:8816–8821
- Tres A, van der Veer G, Perez-Marin MD, van Ruth SM, Garrido-Varo A (2012) Authentication of organic feed by near-infrared spectroscopy combined with chemometrics: a feasibility study. *J Agric Food Chem* 60:8129–8133
- Boulesteix AL (2004) PLS dimension reduction for classification with high-dimensional microarray data. *Stat Appl Genet Mol Biol* 3, Article 33
- Westerhuis JA, De Jong S, Smilde AK (2001) Direct orthogonal signal correction. *Chemometr Intell* 56(1):13–25
- R (Development Core Team R) (2011) *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna
- MacGibbon AHK, Taylor MW (2006) Composition and structure of bovine milk lipids. In: Fox PF, McSweeney PLH (eds) *Advanced dairy chemistry*, 3rd edn. Springer, NY
- Precht D (2001) Cholesterol content in European bovine milk fats. *Nahrung/Food* 45(1):2–8
- Jensen RG (2002) Invited review: the composition of bovine milk lipids: January 1995 to December 2000. *J Dairy Sci* 85:295–350
- Gresti J, Bugaut M, Maniogui C, Bezard J (1993) Composition of molecular species of triacylglycerols in bovine milk fat. *J Dairy Sci* 76:1850–1869
- Dewhurst RJ, Shingfield KJ, Lee MRF, Scollan ND (2006) Increasing the concentrations of beneficial polyunsaturated FAs in milk produced by dairy cows in high-forage systems. *Anim Feed Sci Technol* 131:168–206
- Elgersma A, Tamminga S, Ellen G (2006) Modifying milk composition through forage—a review. *Anim Feed Sci Technol* 131:207–225
- Precht D, Molkentin J (1999) Analysis and seasonal variation of conjugated linoleic acid and further cis/trans-isomers of C18:1 and C18:2 in bovine milk fat. *Kieler Milchwirtsch Forschungsber* 51:63–78
- Capuano E, Boerrigter-Eenling R, van der Veer G, van Ruth S (2013) Analytical authentication of organic products: an overview of markers. *J Sci Food Agric* 93:12–28