Food Research International 136 (2020) 109543

Contents lists available at ScienceDirect

Food Research International

journal homepage: www.elsevier.com/locate/foodres

Opportunities for fraudsters: When would profitable milk adulterations go unnoticed by common, standardized FTIR measurements?

Yuzheng Yang[a,b], Kasper A. Hettinga[a], Sara W. Erasmus[a], Annemieke M. Pustjens[b], Saskia M. van Ruth[a,b,c,*]

[a] Food Quality and Design Group, Wageningen University and Research, P.O. Box 17, 6700 AA Wageningen, the Netherlands
[b] Wageningen Food Safety Research, Part of Wageningen University and Research, P.O. Box 230, 6700 AE Wageningen, the Netherlands
[c] Institute for Global Food Security, Queen's University Belfast, 19 Chlorine Gardens, Belfast BT9 5DL, Northern Ireland, United Kingdom

ARTICLE INFO

Keywords:
Fourier transform infrared
Milk adulteration
Milk composition
Milkoscan measurements
One class classification
Profitability

Chemical compounds studied in this article:
Ammonium chloride (PubChem CID: 25517)
Ammonium sulphate (PubChem CID: 6097028)
Dicyandiamide (PubChem CID: 10005)
Formaldehyde (PubChem CID: 712)
Fructose (PubChem CID: 5984)
Glucose (PubChem CID: 79025)
Hydrogen peroxide (PubChem CID: 784)
Lactose (PubChem CID: 104938)
Maltodextrin (PubChem CID: 68229136)
Melamine (PubChem CID: 7955)
Urea (PubChem CID: 1176)
Sodium bicarbonate (PubChem CID: 516892)
Sodium carbonate (PubChem CID: 10340)
Sodium citrate (PubChem CID: 23666341)
Sodium hydroxide (PubChem CID: 14798)
Starch (PubChem CID: 24836924)
Sucrose (PubChem CID: 5988)

ABSTRACT

Milk is regarded as one of the top food products susceptible to adulteration where its valuable components are specifically identified as high-risk indicators for milk fraud. The current study explores the impact of common milk adulterants on the apparent compositional parameters of milk from the Dutch market as measured by standardized Fourier transform infrared (FTIR) spectroscopy. More precisely, it examines the detectability of these adulterants at various concentration levels using the compositional parameters individually, in a univariate manner, and together in a multivariate approach. In this study we used measured boundaries but also more practical variance-adjusted boundaries to set thresholds for detection of adulteration. The potential economic impact of these adulterations under a milk payment scheme is also evaluated. Twenty-four substances were used to produce various categories of milk adulterations, each at four concentration levels. These substances comprised five protein-rich adulterants, five nitrogen-based adulterants, seven carbohydrate-based adulterants, six preservatives and water, resulting in a set of 360 samples to be analysed. The results showed that the addition of protein-rich adulterants, as well as dicyandiamide and melamine, increased the apparent protein content, while the addition of carbohydrate-based adulterants, whey protein isolate, and skimmed milk powder, increased the apparent lactose content. When considering the compositional parameters univariately, especially protein- and nitrogen-based adulterants did not raise a flag of unusual apparent concentrations at lower concentration levels. Addition of preservatives also went unnoticed. The multivariate approach did not improve the level of detection. Regarding the potential profit of milk adulteration, whey protein and corn starch seem particularly interesting. Combining the artificial inflation of valuable components, the resulting potential profit, and the gaps in detection, it appears that the whey protein isolates deserve particular attention when thinking like a criminal.

1. Introduction

Food adulteration, or food fraud, is described as illegal deception of a food product for economic gain (Spink, Moyer, & Speier-Pero, 2016). Milk is one of the most commonly adulterated foods in the world, while there is currently no decline in the number of posted milk fraud reports (Cavin et al., 2016; Moore, Spink, & Lipp, 2012). Milk adulteration was first documented as adding water to increase volume, but has become

Abbreviations: AC, Ammonium chloride; AR, Arrowroot powder; AS, Ammonium sulphate; BIC, Sodium bicarbonate; CAR, Sodium carbonate; CI, Confidence interval; CIT, Sodium citrate; DIC, Dicyandiamide; FMD, Formaldehyde; FPD, Freezing point depression; FRU, Fructose; FTIR, Fourier transform infrared; GLU, Glucose; HYD, Sodium hydroxide; KNN, K-nearest neighbours; LAC, Lactose; MD, Maltodextrin; MLM, Melamine; OCC, One-class classification; PCA, Principal component analysis; PEA, Pea protein isolate; PX, Hydrogen peroxide; RBF, Radial basis function; SIMCA, Soft independent modelling of class analogies; SMP, Skimmed milk powder; SNF, Solids non-fat; SOY, Soy protein isolate; STA, Corn Starch; SU, Sucrose; SVM, Support vector machine; TS, Total solids; UHT, Ultra-high temperature; URE, Urea; WMP, Whole milk powder; WPI, Whey protein isolate

* Corresponding author at: Food Quality and Design Group, Wageningen University and Research, P.O. Box 17, 6700 AA Wageningen, the Netherlands.
E-mail address: saskia.vanruth@wur.nl (S.M. van Ruth).

https://doi.org/10.1016/j.foodres.2020.109543
Received 16 March 2020; Received in revised form 9 July 2020; Accepted 9 July 2020
Available online 14 July 2020

more sophisticated by using various materials with different purposes (Cassoli, Sartori, Zampar, & Machado, 2011). The pivotal 2008 Chinese milk incident, where infant formula, along with other milk products, was contaminated with melamine, shone a bright light on milk fraud, with its devastating effects still lingering 10 years later (Li, Sijtsema, Kornelis, Liu, & Li, 2019). This scandal not only exposed the malpractices within the dairy industry, but it also demonstrated how fraud affects consumer confidence as the Chinese consumer's preferences shifted towards products with European or New Zealand origin (Kendall et al., 2019). Nowadays, there are countless ways in which milk products are adulterated.

Several categories of substances are known as potential milk adulterants. The nitrogen-rich chemicals (e.g., melamine and ammonium salts) are not easily detected by traditional nitrogen-based protein determination approaches, as they are used to boost the apparent protein content (Finete, Gouvêa, Marques, & Netto, 2013). Carbohydrates can also be fraudulently added to milk to increase the apparent lactose content (Liu, Ren, Liu, & Guo, 2015), while preservatives such as sodium hydroxide and sodium carbonate can be used as a way to prevent or conceal microbial milk spoilage, by means of neutralizing the acidity originating from microbial growth (Singh & Gandhi, 2015; Tripathy, Ghole, Deep, Vanjari, & Singh, 2017). It is known that components extracted from milk are likely to be used for milk adulteration as these types of adulterants are less likely to be detected. For instance, whey protein (a by-product of cheese-making) is a cheaper protein source and has been used to increase the protein content of milk (Liu et al., 2015). To aid in the detection of milk adulterants, advanced analytical approaches have been developed, such as liquid chromatography and mass spectrometry (Abernethy & Higgs, 2013; Luykx et al., 2007; Nascimento, Santos, Pereira-Filho, & Rocha, 2017). However, since these are high-cost, time consuming and labour-intensive as screening methods, it is impractical to use them for testing a large number of samples.

As a result, rapid or high-throughput screening methods for food analyses, which allow very early intervention when anomalies are discovered, have gained increasingly more attention. Fourier transform infrared spectroscopy (FTIR) is one such method. It is a rapid, cost-efficient and user-friendly technique that has primarily been applied for routine quality assessment. Together with multivariate analysis, FTIR has been used in various studies for determining milk adulterations with promising results (Cassoli, Sartori, & Machado, 2011; Coitinho et al., 2017; Jha, Jaiswal, Borah, Gautam, & Srivastava, 2014; Nicolaou, Xu, & Goodacre, 2010). Moreover, automated equipment based on the FTIR has been developed to determine milk properties such as its gross chemical composition (i.e. protein, fat, lactose, total solids, solids non-fat), density, and freezing point depression. These results have been widely utilized to determine the economic value of raw milk deliveries and consequently the earnings of farmers.

Milk payment schemes are generally based on quality and hygiene, while at the same time they have other objectives which includes avoiding adulteration, setting accurate milk prices to reflect its value, etc. (Sneddon, Lopez-Villalobos, Hickson, & Shalloo, 2013). In the Netherlands, the farm gate milk price is based on the farm-specific protein, fat and lactose yield, minus the fixed costs such as transportation expense, plus the premium for special attributes (Royal Friesland Campina, 2019). According to this milk payment scheme, the milk price at farm gate primarily depends on milk composition.

An earlier study has shown that the valuable components of milk are major economic drivers for milk fraud (Yang et al., 2019). The automated equipment based on the FTIR technique is commonly used to determine the composition of milk in order to calculate how much should be paid for the milk (Qlip, 2019). However, the capacity of this technique in detecting common adulterants is not well known, especially when testing for many different categories of adulterants. Thus, the present study aims to explore the impact of common milk adulterants on the apparent protein, fat, and lactose content, and other parameters of UHT milk measured by standardized FTIR, combined with their detectability and the potential economic profit from their addition. Anomaly detection is conducted considering the individual measurement parameters in a univariate manner, but also combined using a multivariate approach. For both approaches, control group data of a set of 15 genuine milk samples are used for comparison to the adulterated samples. Furthermore, since it is acknowledged that these 15 samples do not cover the full range of natural variation in practice, we also worked with more practical variance-adjusted boundaries.

2. Materials and methods

2.1. Reference material

A total of twelve commercial UHT full-fat milk samples from ten different brands were purchased from supermarkets in the Netherlands, during winter (January to February 2018). The sample set included four samples processed in the Netherlands, four processed in Germany, and four processed in Belgium. All samples were stored at room temperature and analysed before their expiry date.

For the adulteration studies, the samples obtained per country were pooled to end up with three pooled samples. Consequently, each pooled sample comprised four milk samples (of the same country) which were mixed at a ratio of 1:1:1:1 w/w. The three milk pools (100 g each) were prepared, to which adulterants were added at different levels. The 15 unadulterated milk samples (12 individual commercial samples and the three milk pools) were used as the control samples in the study.

2.2. Adulterants

Twenty-four adulterants were added in different amounts to the three pooled samples. They were categorized into five groups and listed as follows: (1) protein-rich adulterants including whole milk powder (WMP), skimmed milk powder (SMP), whey protein isolate (WPI), pea protein isolate (PEA) and soy protein isolate (SOY); (2) nitrogen-based adulterants including urea (URE), melamine (MLM), ammonium sulphate (AS), ammonium chloride (AC) and dicyandiamide (DIC); (3) carbohydrate-based adulterants including sucrose (SU), glucose (GLU), corn starch (ST), lactose (LAC), fructose (FRU), maltodextrin (MD) and arrowroot powder (AR); (4) preservatives including sodium citrate (CIT), sodium carbonate (CAR), sodium bicarbonate (BIC), sodium hydroxide (HYD), formaldehyde (FMD) and hydrogen peroxide (PX); (5) water. The nitrogen-based adulterants, preservatives, sucrose, glucose, lactose, fructose maltodextrin and starch from corn were purchased from Sigma Aldrich (St. Louis, MO, USA). The other substances (whole milk powder, skimmed milk powder, whey protein isolate, pea protein isolate, soy protein isolate and arrowroot powder) were purchased from the local suppliers in the Netherlands (January to February 2018), and were stored at room temperature before use. The detailed information of these substances is provided in Table S1 (Supplementary material).

2.3. Adulterations

2.3.1. Single-adulterations

The three milk pools prepared as described in Section 2.1 were spiked with the adulterants according to Eq. (1) to Eq. (3). The weight of the protein-rich adulterants added to the 100 g milk pools was calculated according to Eq. (1):

$$Weight_{protein-rich\ adulterant} = 100\ g \times \left(\frac{Protein_{control} \times a\%}{Protein_{adulterant}} \right) \quad (1)$$

where $Protein_{control}$ is the protein content of the control milk samples, which was 3.5% w/w on average. $Protein_{adulterant}$ is the protein content of the adulterant, while $a\%$ stands for the level of adulteration. The four

levels used for the protein-rich adulterations were 10%, 20%, 30% and 40% w/w (adulterant protein/milk protein content).

The weight of the nitrogen-based adulterants added to the 100 g milk pools was calculated according to Eq. (2):

$$Weight_{nitrogen\ adulterant} = 100\ g \times \left(\frac{Protein_{control} \times a\%}{f \times N_{adulterant}} \right) \quad (2)$$

where $Protein_{control}$is the protein content of the control milk samples, which was 3.5% w/w on average. f is the conversion factor of nitrogen to protein, which is equal to 6.38 for the milk protein. $N_{adulterant}$ is the nitrogen content of the nitrogen-based adulterant. $a\%$ stands for the level of adulteration. The four levels used for the nitrogen-based adulterations were 10%, 20%, 30% and 40% w/w (adulterant protein/milk protein content).

The weight of the carbohydrate-based adulterants added to 100 g milk pool was calculated according to Eq. (3):

$$Weight_{carbohydrate\ adulterant} = 100\ g \times \left(\frac{TS_{control} \times a\%}{TS_{adulterant}} \right) \quad (3)$$

where $TS_{control}$ is the total solids content of the control milk samples, which was 13.0% w/w on average. $TS_{adulterant}$ is the total solids content of the adulterant. $a\%$ stands for the level of adulteration. The four levels used for the carbohydrate-based adulterations were 10%, 20%, 30% and 40% w/w (adulterant TS/milk TS content). The four levels of the preservative contaminants added to the 100 g milk pools were 0.05 g, 0.10 g, 0.15 g and 0.20 g. More detailed information of the single adulteration is provided in Table 1. Six levels were used to test the addition of water to milk: 5%, 10%, 20%, 30%, 40%, and 50% w/w (water/milk). All the samples were stirred for 20 min after the adulterants were added. In total, 294 single-spiked milk samples were thus prepared.

2.3.2. Combined-adulterations

Combined-adulterations were made according to the following steps: 40 g of water was added to 100 g of a milk pooled sample, after which a single adulterant from the protein-rich, nitrogen-based or carbohydrate-based adulterant groups (described in Section 2.2) was added to the diluted pool to increase the apparent protein content with 40% w/w (adulterant protein/milk protein content, for the protein-rich and nitrogen-based adulterations) or to increase the apparent total solids content with 40% w/w (adulterant TS/milk TS content, for the carbohydrate-based adulterations). More detailed information of the combined-adulteration is provided in Table 1. A total of 17 adulterants were spiked to the three milk pools, resulting in a total of 51 combined-adulterated samples. Together with the single adulterated samples, the final adulterant test set comprised a total of 345 adulterated samples. Furthermore, 15 control samples were included, thus resulting in a total of 360 samples to be analysed.

2.4. Measurements

All the samples were measured in duplicate using the MilkoScan FT120 instrument (Foss Electric, Hilleroed, Denmark), with wavenumbers of the spectrum from 5000 to 930 cm^{-1}. The equipment uses the principle of FTIR, and provides a series of milk compositional parameters, such as protein, fat, lactose, total solids (TS), solids non-fat (SNF), density and freezing point depression (FPD). The raw FTIR spectrum could not be obtained due to limitations of the instrument. All the samples were prepared at room temperature and measured within two hours after preparation.

2.5. Statistical analysis

2.5.1. Linear models between the milk pools and their adulterated counterparts

Linear regression was performed between the apparent readings of the three milk pools and their adulterated counterparts at the four adulterated levels for each compositional parameter. The average values of the three pools for the slope (m) and R-square (R^2) were calculated and presented.

Table 1
The amount of adulterants added into the three milk pools (per 100 g) for the single adulteration and combined adulteration.

Adulterant category	Adulterant	Single adulteration				Combined adulteration	
		Level 1 (g)	Level 2 (g)	Level 3 (g)	Level 4 (g)	Adutlerant (g)	Water (g)
Protein-rich adulterants	WMP	1.36	2.72	4.09	5.45	13.07	40.00
	SMP	0.99	1.98	2.97	3.95	9.49	40.00
	WPI	0.38	0.75	1.13	1.51	3.61	40.00
	SOY	0.39	0.78	1.17	1.56	3.73	40.00
	PEA	0.43	0.85	1.28	1.71	4.10	40.00
Nitrogen-based adulterants	URE	0.12	0.24	0.35	0.47	1.13	40.00
	MLM	0.08	0.16	0.25	0.33	0.79	40.00
	AC	0.21	0.42	0.63	0.84	2.01	40.00
	AS	0.26	0.52	0.78	1.04	2.49	40.00
	DIC	0.08	0.16	0.25	0.33	0.79	40.00
Carbohydrate-based adulterants	SU	1.30	2.60	3.90	5.20	12.48	40.00
	GLU	1.30	2.60	3.90	5.20	12.48	40.00
	FRU	1.30	2.60	3.90	5.20	12.48	40.00
	LAC	1.30	2.60	3.90	5.20	12.48	40.00
	MD	1.30	2.60	3.90	5.20	12.48	40.00
	STA	1.30	2.60	3.90	5.20	12.48	40.00
	AR	1.53	3.06	4.59	6.12	14.68	40.00
Preservatives	CIT	0.05	0.10	0.15	0.20	N.A.	N.A.
	CAR	0.05	0.10	0.15	0.20	N.A.	N.A.
	BIC	0.05	0.10	0.15	0.20	N.A.	N.A.
	FMD	0.05	0.10	0.15	0.20	N.A.	N.A.
	PX	0.05	0.10	0.15	0.20	N.A.	N.A.
	HYD	0.05	0.10	0.15	0.20	N.A.	N.A.

AC: ammonium chloride; AR: arrowroot powder; AS: ammonium sulphate; BIC: sodium bicarbonate; CAR: sodium carbonate; CIT: sodium citrate; DIC: dicyandiamide; FMD: formaldehyde; FRU: fructose; GLU: glucose; HYD: Sodium hydroxide; LAC: lactose; MD: maltodextrin; MLM: melamine; PEA: pea protein isolate; PX: hydrogen peroxide; SMP: skimmed milk powder; SOY: soy protein isolate; STA: starch; SU: sucrose; URE: urea; WMP: whole milk powder; WPI: whey protein isolate.

2.5.2. Univariate analysis: Determination of boundaries for each compositional parameter

The mean value and standard deviation for each compositional parameterwere calculated based on the 15 control samples. The values of the 0.5th and 99.5th percentile for the seven compositional parameters (i.e. protein, fat, lactose, TS, SNF, FPD and density) were used as the measured boundaries for the determination of the measured samples.

Based on a programme concerning the measurement of Dutch raw milk samples for legislatorial control (Zuivelverordening, 2000), the variance of over 3 million milk samples has been documented. The standard deviation (SD) for each compositional parameterin the above-mentioned dataset is roughly double the values of the dataset of control samples in our study. To expand the variation of the control samples in a practicable way, the data of the 15 control samples was converted into a variance-adjusted dataset. For this dataset, the mean values remained unchanged, but the SD values were adjusted to twice the value of the SD of the measured dataset for each compositional parameterusing Eq. (4):

$$X_{adjusted} = \left(\frac{X-\mu}{\sigma} \times 2\sigma\right) + \mu \quad (4)$$

where $X_{adjusted}$ is the data for the variance-adjusted set, X is the measured data of the sample, μ is the mean value of the 15 control samples for each compositional parameter, and σ is the SD of the 15 control samples for each compositional parameter. Next to that, the values of 0.5th percentile and 99.5th percentile for the variance-adjusted dataset for the seven compositional parameters were calculated, and used as the variance-adjusted boundaries for the determination of suspect samples.

Both the measured boundaries and the variance-adjusted boundaries were then applied to the adulterant test set. Samples with any compositional parameter exceeding the boundaries were labelled as "suspected adulterations". The two datasets are presented in Table S2 and S3 (Supplementary material).

2.5.3. Multivariate analysis: Determination of threshold for milk with one class classification models

Principal component analysis (PCA) of the data acquired from the MilkoScan measurements was first performed, using the pre-processing method of auto-scaling, to explore the presence of clustering for the different groups of adulterated samples. In terms of determining the threshold for the control samples with the multivariate analysis, three one class classification (OCC) models were calculated, namely, soft independent modelling of class analogies (SIMCA), k-nearest neighbours (KNN), and support vector machine (SVM). SIMCA is based on PCA. It computes PC models for the representative class and classifies the unknown samples. SIMCA focuses more on the similarities among samples within a class and thus is widely used for OCC models (Gurbanov, Gozen, & Severcan, 2018). The performance of a SIMCA model depends on the number of selected factors (n). KNN evaluates the distance from the object to its k nearest neighbours, where the model performance depends on the k value. KNN requires no prior knowledge about the data distribution, it is robust to a noisy way of training data and is suitable for small training sets (Beebe, Pell, & Seasholtz, 1998). SVM is another suitable approach for a dataset with a limited number of training samples (Gholami & Fakhari, 2017). SVM evaluates the distance from the object to the boundary of the model, based on the Gaussian radial basis function (RBF) kernel, which is defined according to Eq. (5),

$$k(X1, X2) = Exp(-\gamma ||X1 - X2||^2) \quad (5)$$

Accordingly, a SVM model performance depends on the model parameterγ.

The measured dataset ($n = 15$) and the variance-adjusted dataset ($n = 15$) of the control samples were used separately as training sets for the development of classification models. The training set was subjected to leave-30%-out with random cross-validation with 100 repetitions. Autoscaling was applied to the dataset in conjunction with the three classifiers. A significance level of 1% ($p = 0.01$) was used to determine the critical classification limit. The adulterant test set, comprising 345 adulterated milk samples, was then subjected to the developed models. The adulterant test set comprised five sub-sets, i.e. protein-rich adulterated sub-set ($n = 75$), nitrogen-based adulterated sub-set ($n = 75$), carbohydrate-based adulterated sub-set ($n = 105$), preservative contaminated sub-set ($n = 72$), and water diluted sub-set ($n = 18$). The three OCC models were estimated by applying the following model parameters: the number of factors n for SIMCA was selected from consecutive number 1–7; k for the KNN model was selected for the consecutive numbers 1–10; γ for the SVM was selected from 10^{-9}, 10^{-8}, 10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1} and 1. The balanced accuracy approach was used to evaluate the overall performance of the models due to its skewed class distribution (García, Mollineda, & Sánchez, 2009; Sokolova, Japkowicz, & Szpakowicz, 2006). The balanced accuracy was calculated according to Eq. (6):

$$Balanced\ accuracy = \frac{(True\ positive\ rate + True\ negative\ rate)}{2} \quad (6)$$

The optimal model parameter for the best performing model was selected accordingly.

All the statistical analyses in this study were performed using R 3.6.1 software (R Foundation for Statistical Computing, Vienna, Austria).

2.6. Potential profit calculations

The milk payment scheme from Royal Friesland Campina, the largest dairy company in the Netherlands (Royal Friesland Campina, 2019), was used to calculate the financial implication of the milk compositional changes caused by common adulterants. The net profit changes per 100 kg of milk caused by the adulterants were calculated using Eq. (7):

$$Net\ profit\ change = Increased\ profit - Cost\ of\ adulterant \quad (7)$$

The prices of the substances specified in the United States Department of Agriculture National Agricultural Statistics Service database (USDA-NASS, 2019) were applied to calculate the *Cost of adulterant*. According to the payment scheme, the *Increased profit* depended on the increased protein, fat and lactose contents for 100 kg milk, and was calculated using Eq. (8):

$$Increased\ profit = 1000 \times \sum_i \Delta_i \times p_i \quad (8)$$

where i stands for the payment milk composition, i.e. protein, fat and lactose, Δ_i stands for the concentration difference (g/100 g) between the adulterated and control samples, p_i stands for the price of the protein, fat or lactose, which changes every month. According to the milk payment scheme for the last year from October 2018 to September 2019, the average p values for protein, fat and lactose were 590, 295, 59 euro per 100 kg, respectively.

3. Result and discussion

3.1. Natural variation of the control samples

Variation in composition of the control milk samples was observed (Table 2; Supplementary material Table S2). The 99% confidence interval for the milk protein content ranged from 3.28% to 3.81% w/w. More specifically, the average protein contents of the milk samples from the Netherlands, Germany and Belgium (i.e. the three pools) were 3.73%, 3.55% and 3.50% w/w, respectively (data not shown), which are in agreement with the natural variation reported for milk produced in these countries (Heck, van Valenberg, Dijkstra, & van Hooijdonk,

Table 2
The mean values and standard deviation (SD) of the composition of the control milk samples (n = 15), and the boundaries based on the measured dataset and variance-adjusted dataset.

Parameter		Protein (% w/w)	Fat (% w/w)	Lactose (% w/w)	FPD (°C)	TS (% w/w)	SNF (% w/w)	Density (g/L)
Mean		3.57	3.77	4.67	0.564	13.61	9.81	1036
Standard deviation		0.15	0.11	0.06	0.012	0.23	0.22	1
Measured boundary	Lower boundary	3.28	3.63	4.57	0.588	13.17	9.38	1034
	Upper boundary	3.81	4.01	4.77	0.548	13.97	10.19	1037
Variance-adjusted boundary	Lower boundary	2.98	3.49	4.48	0.531	12.73	8.94	1031
	Upper boundary	4.06	4.26	4.87	0.611	14.34	10.58	1038

FPD: Freezing point depression; TS: Total solids; SNF Solids non-fat.

2009; Rattray & Jelen, 1996). The 99% confidence interval of the fat content ranges from 3.63% to 4.01% w/w, while the variation in lactose content appears small, ranging from 4.57% to 4.77% w/w. It is known that the variation of milk composition can be affected by a series of external factors such as season, origin, cow breed, etc. The variation of the composition of the UHT milk samples in the current study was generally small, which is mainly due to the standardization steps during processing normalized the milk composition. To expand the variation of milk composition, the measured dataset was converted to a variance-adjusted dataset to allow larger natural variation, as previously described in Section 2.5.2. Both datasets were used to evaluate the potential to detect suspected milk adulterations.

3.2. The impact of adulterants on the payment parameters

The impact of each adulterant on the apparent milk composition is presented in Table 3. The protein-rich and carbohydrate-based adulterant resulted in an increase in the apparent protein content and apparent lactose content, respectively. It is noted that most of the nitrogen-based adulterations resulted in no significant increase in apparent protein concentration. To gain insight into the impact of each adulterant on all parameters, the PCA results of single adulterated samples were presented (Fig. 1). The different adulterants showed partly overlapping results in the PCA plot, with larger differences for the highest adulteration levels. Among the different adulterants, the nitrogen-based adulterants and the water-diluted samples showed clear distinction from the other categories, whereas the protein-rich and carbohydrate-based adulterants overlap to some extent.

3.2.1. Impact of the protein-rich adulterations

The protein-rich adulterants contain protein, fat and lactose, which are all related to the milk price. Consequently, it is important to determine the changes of these compositional parameters after adulterations. In Table 3, it shows that the addition of protein-rich adulterants not only increased the apparent protein concentrations, but it also increased it to different extents. The WPI adulteration resulted in the largest increase of apparent protein (slope = 0.33), followed by the WMP and SMP (both slopes equal to 0.28) (Table 3). The samples spiked with soy and pea protein showed a less apparent protein increase when compared to the milk-based adulterants, which may be due to the

Table 3
The linear model performance comparing apparent results and the spiked levels, for the parameters of protein, fat, lactose and FPD for each adulterant. The outcome of the total solids, solids non-fat and density are shown in Table S4 (Supplementary material).

Category	Adulterant	Protein		Fat		Lactose		FPD	
		m	R^2	m	R^2	m	R^2	m	R^2
Protein-rich adulterants	WMP	0.28	1.00	0.27	0.99	0.41	1.00	0.05	1.00
	SMP	0.28	0.99	−0.05	0.84	0.45	0.99	0.05	1.00
	WPI	0.33	1.00	−0.04	0.67	−0.02	0.36	0.02	0.97
	SOY	0.21	0.99	−0.03	0.56	0.00	0.38	0.01	0.86
	PEA	0.27	0.98	−0.05	0.77	−0.02	0.28	0.00	0.49
Nitrogen-based adulterants	URE	−0.08	0.97	−0.12	0.93	−0.04	0.88	−0.01	0.95
	MLM	0.08	0.90	−0.04	0.60	−0.02	0.55	0.01	0.88
	AC	−0.18	0.99	−0.19	0.97	−0.12	0.99	0.01	0.71
	AS	−0.18	1.00	−0.20	0.98	0.10	0.97	−0.18	1.00
	DIC	0.17	0.97	−0.03	0.50	0.00	0.44	0.01	0.65
Carbohydrate-based adulterants	SU	−0.03	0.68	−0.05	0.69	1.13	1.00	0.22	1.00
	GLU	−0.01	0.61	0.00	0.06	1.09	1.00	0.20	1.00
	LAC	−0.02	0.86	−0.06	0.83	1.09	1.00	0.09	0.98
	FRU	−0.03	0.90	−0.04	0.78	0.82	0.99	−0.01	0.89
	MD	−0.04	0.94	−0.09	0.92	1.19	0.99	0.37	1.00
	STA	0.06	0.97	0.51	0.96	0.34	0.95	0.15	0.98
	AR	0.07	0.98	0.61	1.00	0.42	0.99	0.18	1.00
Preservatives	CIT	0.03	0.81	−0.03	0.34	−0.01	0.25	0.00	0.12
	CAR	0.02	0.66	−0.04	0.82	0.03	0.73	0.01	0.72
	BIC	0.01	0.61	−0.02	0.39	0.01	0.39	0.01	0.83
	FMD	0.02	0.29	−0.04	0.71	0.06	0.94	0.02	0.98
	PX	0.03	0.50	−0.02	0.15	0.02	0.43	0.01	0.97
	HYD	0.00	0.02	−0.08	0.98	0.05	0.72	0.01	0.57

The slope (m) and R-squared (R2) values are the average values for the linear regressions of three pools and their adulterated counterparts.
AC: ammonium chloride; AR: arrowroot powder; AS: ammonium sulphate; BIC: sodium bicarbonate; CAR: sodium carbonate; CIT: sodium citrate; DIC: dicyandiamide; FMD: formaldehyde; FPD: freezing point depression; FRU: fructose; GLU: glucose; HYD: Sodium hydroxide; LAC: lactose; MD: maltodextrin; MLM: melamine; PEA: pea protein isolate; PX: hydrogen peroxide; SMP: skimmed milk powder; SOY: soy protein isolate; STA: starch; SU: sucrose; URE: urea; WMP: whole milk powder; WPI: whey protein isolate.

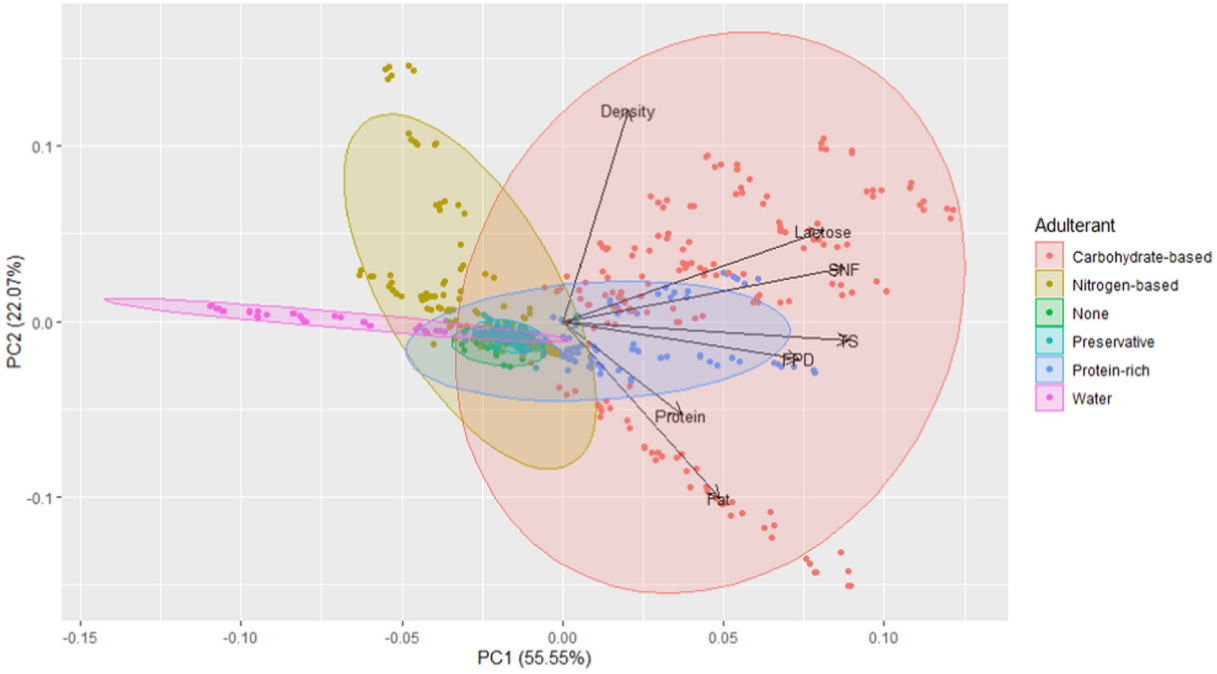

Fig. 1. The Principal Component Analysis (PCA) scores bi-plot of the first two PC dimensions of all milk samples based on the data obtained from the MilkoScan measurements. For interpretation of the different colours, refer to the web version of the paper.

different properties of the proteins in these protein-rich substances. The milk proteins, which consist of mainly casein and whey protein, are likely the main contributors to the change of the apparent protein concentration for the WMP, SMP and WPI adulterated samples. Moreover, the commercial soy and pea protein isolates can easily form incompletely dissolved suspensions during dissolution (Zhang, Liang, Tian, Chen, & Subirade, 2012). In fact, this phenomenon was observed during the sample preparation, and could likely have influenced the results.

Not only the apparent protein content of the samples increased when the protein-rich adulterants were added, similarly the fat and lactose concentrations of the samples also increased. This is likely due to the complex composition of the adulterants, i.e. there are non-protein components in the presence. For example, Table 3 shows that the WMP and SMP adulterations increase the apparent lactose concentration, which is attributable to the fact that the carbohydrate contents of the WMP and SMP (36.5% and 50.5% w/w, respectively) are much higher than those of the other adulterants (< 3% w/w) (Table S1, Supplementary material). It is noted that all the protein-rich adulterants changed the apparent protein/fat and protein/lactose ratio. Consequently, by using multivariate analysis, it may be possible to identify the adulterated milk.

Overall, due to the variation in composition of the protein-rich adulterants, the apparent protein, fat and lactose contents of the adulterated samples were affected differently. The WPI increased the apparent protein content of the corresponding adulterated milk samples the most, whereas the WMP increased the apparent protein, fat and lactose content of the corresponding adulterated milk samples. The soy and pea protein adulteration increased apparent protein contents to a lesser extent.

3.2.2. Impact of the nitrogen-based adulterations

The international reference method for milk protein determination, the Kjeldahl method (International Organization for Standardization [ISO], 2014), is based on the nitrogen content of a sample. The nitrogen-based adulterations mainly aim at increasing the apparent protein content of milk samples, through increasing the nitrogen content. Therefore, the change of apparent protein content after these adulterations is further discussed in this section.

As shown in Table 3 under the nitrogen-based adulterants category, the apparent protein concentration of the samples spiked with DIC and MLM was positively related with the adulteration level, i.e. the more adulterant added, the higher the apparent protein content. On the contrary, the apparent protein content of the samples spiked with URE, AS and AC showed the opposite trend - the more of these adulterants added, the lower the apparent protein content. This is likely related to the spectral features of the adulterants. For instance previous study reported that MLM generated several absorption bands in FTIR spectrum (Jawaid, Talpur, Sherazi, Nizamani, & Khaskheli, 2013), which may relate to the increase of the apparent protein content. Considering AS and AC have no specific carbon atom bonds, it is assumed that they produced no specific bands in the FTIR spectrum related to the protein structure to be identified.

To conclude, the addition of the DIC and MLM increased the apparent protein content, whereas the other nitrogen-based adulteration showed no, or an opposite trend. It is worth noting that, when comparing with the traditional method, namely Kjeldahl method, the impact of nitrogen rich adulterants on determination of protein content is smaller for FTIR measurement.

3.2.3. Impact of the carbohydrate-based adulterations

The carbohydrate-based adulterations aim to change the apparent lactose or dry matter content of milk. From the group of carbohydrate-based adulterants of Table 3, we can see that the MD adulteration resulted in the largest slope (1.19) for apparent lactose concentration, followed by the SU (1.13), LAC (1.09) and GLU (1.09) adulteration. The apparent lactose concentration of the STA and AR adulterations increased to a lesser extent compared to the others. However, it is important to note that these two starch adulterants also caused an increase in the apparent fat concentration. Currently, it is not known what might have caused this response, as it has also not been reported in literature before. The sugars used in this study share the same chemical bonds, such as C-C stretching modes and C-O-H bending modes, which likely resulted in similar absorption bands in FTIR spectrum (Bureau et al., 2009; Kačuráková & Mathlouthi, 1996; Kačuráková & Wilson, 2001). This can possibly explain the increase in apparent lactose reading

observed for the carbohydrate adulterations.

To summarize, the addition of MD increased the apparent lactose content the most, followed by SU, LAC and GLU. The addition of STA and AR increased both apparent lactose and fat content.

3.2.4. Impact of the preservatives addition

In the PCA plot (Fig. 1), it is seen that the samples spiked with the preservatives overlapped with the control samples. This overlap indicates that the two groups are very similar and that it may be difficult to distinguish them. Generally, the addition of preservatives can increase the FPD by increasing the number of ion particles in the solution, which usually can be discovered by conductivity measurement. The lack of conductivity sensor limited the performance of FTIR analysis. Consequently, the result of the FTIR measurements showed no change in the FPD after the preservatives were added to the milk samples, while the changes of other compositional parameters were also very small (Table 3).

3.3. When will the adulterations be detected?

3.3.1. Univariate detection by apparent concentrations of individual components/features

To explore at which concentrations the milk adulterations would raise a flag because of an unusual compositional trait, the adulterated samples were held against the boundaries from the measured dataset and the variance-adjusted dataset (Fig. 2). Using the boundaries of the measured dataset, all the protein-rich, nitrogen-based and carbohydrate-based adulterations would be flagged (red) except for the lowest level of melamine adulteration of pool C (green) (Fig. 2a). Note that four compositional values of control samples out of 105 measurements exceeded the upper boundaries as well. They concern one milk sample (control sample 4) that exceeds the boundaries for protein, solids and solids non-fat contents, and one other sample (control sample 9) that shows an extraordinary high protein content. It is most likely that these two milk samples are of exceptional compositional quality rather than that they are adulterated. Since a large variance is faced in practice, the variance-adjusted boundaries, which are less strict but more practical, were considered as well (as described in Section 2.5.2). Based on these variance-adjusted boundaries, all the control samples were within the boundaries (data not shown). The detailed results of the adulterated samples are shown in Fig. 2b. It shows that the carbohydrate adulterations were all flagged, and higher concentrations (level 3 and 4) of protein-rich and nitrogen-based adulterations were flagged too. In addition, all the combined adulterations were flagged due to their unusual compositional features. However, the addition of protein-rich and nitrogen-based adulterants at lower levels, and the addition of preservatives, would not stand out. It can be summarised that, with the application of variance-adjusted boundaries, due to unusual apparent compositional features, carbohydrate adulterations, and higher concentration of protein-rich and nitrogen-based adulterations would be detected but inflation of the apparent protein content by 10–20% with WPI, SOY, PEA, MLM, AC and DIC as well as the addition of most preservatives would most likely go unnoticed. Fraudsters may benefit from this gap in detectability of adulteration.

3.3.2. Multivariate detection by combined apparent concentrations of individual components/features

Three OCC models, namely SIMCA, KNN, and SVM, have been developed separately according to the measured and variance-adjusted datasets. The results of the three OCC models for the two scenarios (i.e. based on the measured dataset and the variance-adjusted dataset) were compared. It appears that the SIMCA model achieved the best

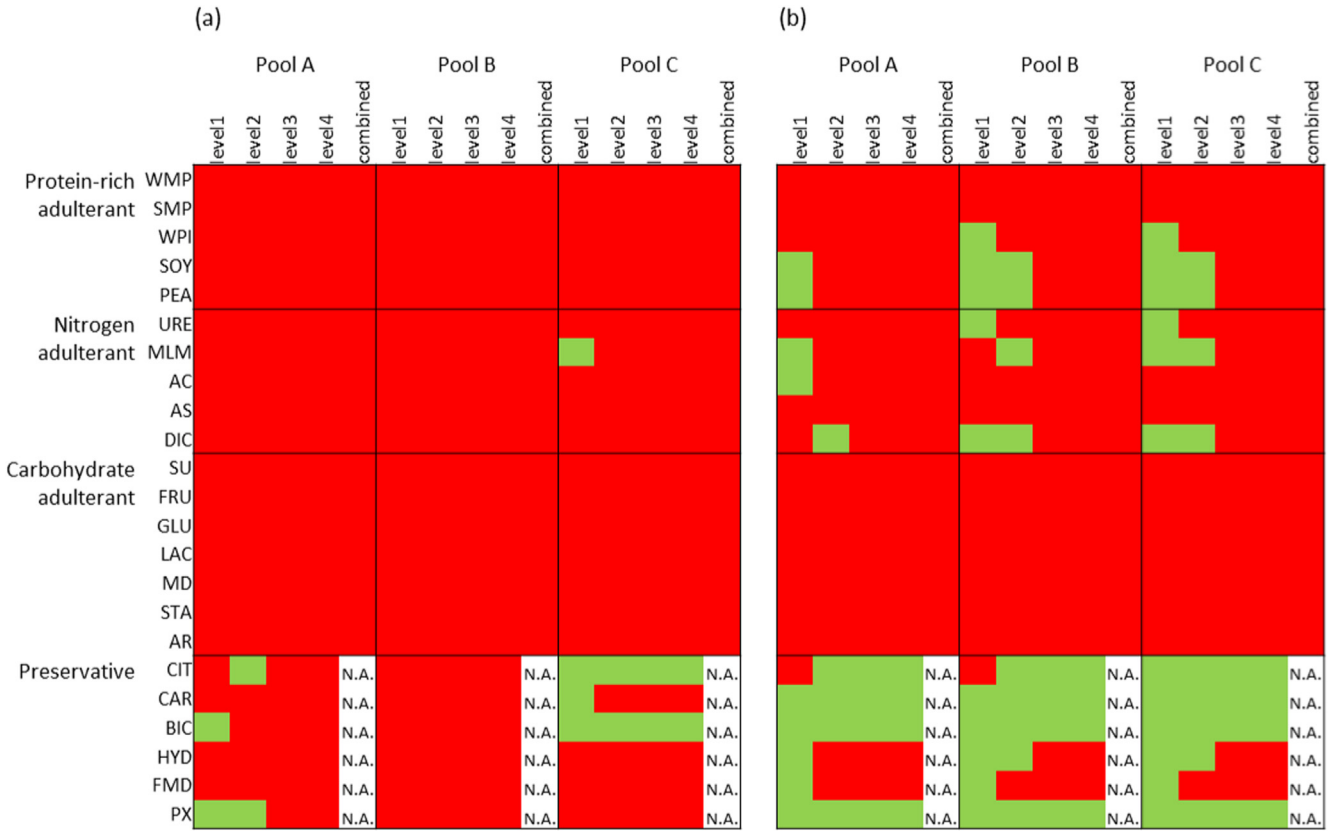

Fig. 2. The results of the adulterant test sets for the three milk pools based on (a) the measured boundaries; and (b) the variance-adjusted boundaries, indicating the potential to identify suspected milk adulterations. The samples with all results within the boundaries are coloured green, while the rest are coloured red. The full names of the adulterant are shown in the abbreviation list. For interpretation of the different colours, refer to the web version of the paper.

Table 4
The results of the one-class classification models developed from the measured dataset and the variance-adjusted dataset. All the present values are the average values of 100 repetitions of cross validation, and the values stand for the accuracy of the corresponding datasets (i.e. true positive for the cross validation set, and true negative for the adulterant test set).

Models	Datasets	SIMCA [a] (%)	KNN [b] (%)	SVM [c] (%)
Models based on the measured dataset	Cross validation set	92	92	87
	Adulterant test set	85	78	80
	Balanced accuracy	88	85	83
Models based on the variance-adjusted dataset	Cross validation set	88	97	83
	Adulterant test set	69	57	71
	Balanced accuracy	78	77	77

[a] SIMCA stands for soft independent modelling of class analogies, the SIMCA model with best performance was estimated with number of the factors n = 3 (based on both measured dataset and variance-adjusted dataset).
[b] KNN stands for k-nearest neighbours, the KNN model with best performance was estimated with k = 4 (based on measured dataset), and k = 7 (based on variance-adjusted dataset).
[c] SVM stands for support vector machine, the SVM model with best performance was estimated withγ = 0.0001 (based on measured dataset), and γ = 0.1 (based on variance-adjusted dataset).

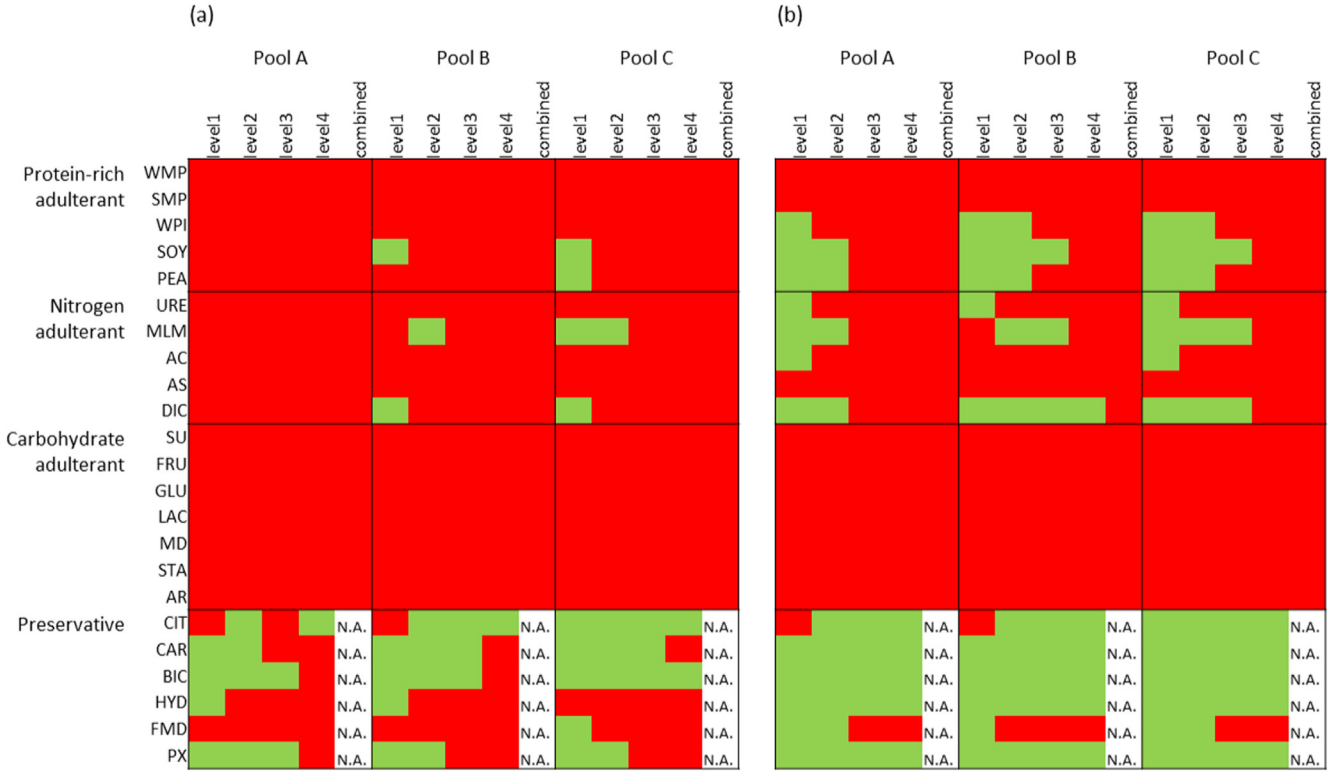

Fig. 3. The results of the adulterant test sets for the three milk pools based on the threshold of the of soft independent modelling of class analogies (SIMCA) model developed from (a) the measured dataset; and (b) the variance-adjusted dataset, indicating the classification of milk adulterations. The samples with results within the threshold are coloured green, while the rest are coloured red. The full names of the adulterant are shown in the abbreviation list. For interpretation of the different colours, refer to the web version of the paper.

performance in both scenarios (Table 4), hence, more detailed results are presented for SIMCA model (Fig. 3). All the control samples were correctly classified by the SIMCA model. It is noted that the model developed from the measured dataset performed better than the model developed from the variance-adjusted dataset (Table 4), similar to the univariate detection. Although an OCC model integrates all aspects and requires only one analysis, it appears that even the best performing model did not result in enhanced detection of adulteration. For instance, considering the measured data set, with univariate comparisons only in one case an adulterant/concentration level combination for all protein-rich, nitrogen-based and carbohydrate-based adulterants/concentration level combinations is not flagged (Fig. 2). However, with the multivariate comparison, eight cases are missed (Fig. 3). Similarly for the variance-adjusted set, 24 are missed using the univariate approach (Fig. 2) and 40 using the multivariate approach (Fig. 3). Obviously, sensitivity may improve with larger numbers of samples, but currently the detection of adulteration seems to work better when using individual compositional parameters.

3.4. The potential profit of the adulterations

Milk adulteration is performed for economic gain, therefore, the economic influence as a result of the change in the composition of milk due to the use of different adulterants was explored. The effect on the net profits generated by the protein-rich, nitrogen-based and carbohydrate-based adulterants was calculated (Table 5). Under the current payment scheme of the raw milk, which is based on yield (in kg) of fat, protein, and lactose, milk dilution does not increase the total price of

Table 5
The economic profit of the minimum and maximum adulteration levels (1 and 4) of the UHT milk for the protein-rich, nitrogen-based and carbohydrate-based adulterants.

Category	Adulterant	Cost (Euro/kg)	Profit increased (Euro/100 kg)		Net profit (Euro/100 kg)	
			Level 1	Level 4	Level 1	Level 4
Protein-rich adulterants	WMP	3.6	3.0	11.0	−1.6	−7.3
	SMP	2.3	2.1	7.4	0.0	−1.1
	WPI	5.0 [a]	2.2	7.5	0.4 [a]	0.3 [a]
	SOY	N.A. [b]	0.7	4.7	–	–
	PEA	N.A. [b]	1.5	6.0	–	–
Nitrogen-based adulterants	URE	0.6	−0.9	−3.3	−1.0	−3.5
	MLM	N.A. [b]	−0.2	1.2	–	–
	AC	N.A. [b]	−2.0	−6.8	–	–
	AS	0.5	−1.8	−6.3	−1.9	−6.9
	DIC	N.A. [b]	1.1	3.8	–	–
Carbohydrate-based adulterants	SU	1.2	−0.2	1.2	−2.3	−7.0
	GLU	0.8	0.6	2.4	−0.4	−1.7
	FRU	0.7	0.3	0.7	−0.6	−2.9
	LAC	0.6	0.5	1.5	−0.2	−1.5
	MD	N.A. [b]	0.3	0.9	–	–
	STA	0.3	1.7	8.0	1.3	6.4
	AR	N.A. [b]	2.5	9.7	–	–

AC: ammonium chloride; AR: arrowroot powder; AS: ammonium sulphate; DIC: dicyandiamide; FMD: formaldehyde; FRU: fructose; GLU: glucose; LAC: lactose; MD: maltodextrin; MLM: melamine; PEA: pea protein isolate; PX: hydrogen peroxide; SMP: skimmed milk powder; SOY: soy protein isolate; STA: starch; SU: sucrose; URE: urea; WMP: whole milk powder; WPI: whey protein isolate.

[a] The cost of whey protein concentrate 34 (based on unit of protein) was used to calculate the price of WPI.

[b] N.A. means no information was found in the USDA NASS database for the cost of adulterants.

milk, and therefore the economic implication due to water dilution was not included in this section.

As shown in Table 5, although the protein-rich milk adulterations increased the gross profit, the net profits of these additions varied, due to the different costs of the raw materials. The WMP and SMP adulterations resulted in a negative net profit, while the WPI adulteration, in contrast, led to a positive net profit. Unlike the WMP and SMP, whey protein, which is the basis of WPI, is a relatively cheaper milk protein, and has been widely considered as a common milk adulterant (De La Fuente & Juárez, 2005; Kartheek, Smith, Muthu, & Manavalan, 2011; Santos, Pereira-Filho, & Rodriguez-Saona, 2013). Overall, the results show that the addition of whey protein into milk is profitable under the current payment scheme, while it is valueless to add WMP and SMP for milk fraud.

Among the nitrogen-based adulterants, the melamine and dicyandiamide adulteration can increase the gross profit, by means of increasing the apparent protein content. On the contrary, the urea, ammonium chloride and ammonium sulphate adulterations decreased the net profit, due to their negative effects on the apparent protein and fat contents (Table 3 and 5). The nitrogen-based adulterants are generally used as fertilizers or other non-edible industrial materials. Their prices are relatively low, and they can usually cause health problems if consumed (Mecker et al., 2012; Moyer, DeVries, & Spink, 2017). The net profitability of all nitrogen-based milk adulterations is negative, according to the available adulterant price in the USDA NASS database. However, due to its positive gross profit, attention should be given to the dicyandiamide milk adulteration considering that fraudsters may have access to cheaper sources or face a different type of payment scheme.

The carbohydrate-based milk adulterations aim to increase apparent lactose content and consequently make profit. Due to the current milk payment scheme, the lactose price is rather low compared to the price for protein and fat (the price ratio of protein, fat and lactose is 10:5:1). Since the addition of carbohydrates such as sucrose, fructose and glucose, increased only the apparent lactose content, the net profit gained through these adulterations were very low, while some were even negative. On the contrary, the starch adulteration increased the apparent protein, fat and lactose content, all of which contributed to extra profit. Ultimately, among all the carbohydrate-based adulterants used, only the starch adulteration is profitable in practice for milk fraud under the Dutch payment scheme.

Overall, based on the prices in the USDA NASS database and the applied Dutch payment scheme calculations, only the whey protein isolate and starch adulterations resulted in positive net profit. The other adulterants were either too costly and/or unable to increase the predicted payment. However, more price-efficient sources of adulterants or other payment criteria may result in another cost-benefit balance. For instance, under payment scheme where milk price is based on the total volume, which may occur in unorganized sector, the effect of dilution could be compensated by adding cheaper milk components or other chemicals, which would make some of the other adulterants tempting to potential criminals.

3.5. Will the profitable adulterations be detected?

As discussed in Section 3.4, protein-rich and nitrogen-based adulterants and starch adulterations, are most likely to lead to a positive net profit. Considering gaps in detection for cases with 10–20% apparent protein content inflation as discussed in Section 3.3, the protein-rich and nitrogen-based adulterants are of key concern. With testing for these payment parameters only, the first type of adulterations will most likely go unnoticed. On the other hand, adulterations with starch will be flagged very quickly. The multivariate approach considering all compositional parameters together did not improve detection. The specificity of the OCC models may be increased by stricter thresholds, although this also poses a higher risk of decreasing the model's sensitivity. It is therefore a trade-off, where the proper threshold should be selected based on the users' demand in practice. Using the full set of spectral data, which was not the subject of this study, may possibly provide relief. However, these raw data are generally not accessible by the routine user, and these automated systems are very much black boxes. Some instruments have a screening model targeted at the presence of specific adulterants, or a 'broad anomaly detection' option which presumably uses whole spectra information, but to what extent such an anomaly model picks up particular adulterations is unknown. The development of the univariate and multivariate approaches in this study was a demonstration for milk authentication using the results of milk composition from routine analysis. The proposed approaches can discover most of common milk adulterations, hence is promising to be applied for fraud mitigation and controls. Of course, more elaborate procedures for detection of individual or groups of adulterants are widely available (ISO, 2009, ISO, 2014; ISO, 2015; AOAC, 2019), but they require time and resources.

CRediT authorship contribution statement

Yuzheng Yang: Conceptualization, Investigation, Formal analysis, Writing - original draft, Writing - review & editing. **Kasper A. Hettinga:** Methodology, Writing - original draft, Writing - review & editing. **Sara W. Erasmus:** Writing - original draft, Writing - review & editing. **Annemieke M. Pustjens:** Writing - review & editing. **Saskia M. Ruth:** Conceptualization, Methodology, Formal analysis, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial

interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The first author received a PhD scholarship from the Sino-Dutch Dairy Development Centre (SDDDC). The authors thank Frans Lettink from the Food Quality and Design group of Wageningen University and Research for the technical assistance regarding the MilkoScan measurements.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodres.2020.109543.

References

Abernethy, G., & Higgs, K. (2013). Rapid detection of economic adulterants in fresh milk by liquid chromatography–tandem mass spectrometry. *Journal of Chromatography A, 1288*, 10–20. https://doi.org/10.1016/j.chroma.2013.02.022.

AOAC. (2019). AOAC SMPR 2012.002: Standard Method Performance Requirements for Whey Protein: Casein Ratio in Infant Formula. Journal of AOAC International, 96(3), 501-501. doi:10.5740/jaoac.int.SMPR2012.002.

Beebe, K. R., Pell, R. J., & Seasholtz, M. B. (1998). *Chemometrics: a practical guide, Vol. 4*. New York: Wiley-Interscience.

Bureau, S., Ruiz, D., Reich, M., Gouble, B., Bertrand, D., Audergon, J.-M., & Renard, C. M. G. C. (2009). Application of ATR-FTIR for a rapid and simultaneous determination of sugars and organic acids in apricot fruit. *Food chemistry, 115*(3), 1133–1140. https://doi.org/10.1016/j.foodchem.2008.12.100.

Cassoli, L. D., Sartori, B., & Machado, P. F. (2011). The use of the Fourier Transform Infrared spectroscopy to determine adulterants in raw milk. *Revista Brasileira de Zootecnia, 40*(11), 2591–2596.

Cassoli, L. D., Sartori, B., Zampar, A., & Machado, P. F. (2011). An assessment of Fourier Transform Infrared spectroscopy to identify adulterated raw milk in Brazil. *International Journal of Dairy Technology, 64*(4), 480–485. https://doi.org/10.1111/j.1471-0307.2011.00711.x.

Cavin, C., Cottenet, G., Blancpain, C., Bessaire, T., Frank, N., & Zbinden, P. (2016). Food Adulteration: From Vulnerability Assessment to New Analytical Solutions. *CHIMIA International Journal for Chemistry, 70*(5), 329–333. https://doi.org/10.2533/chimia.2016.329.

Coitinho, T. B., Cassoli, L. D., Cerqueira, P. H. R., da Silva, H. K., Coitinho, J. B., & Machado, P. F. (2017). Adulteration identification in raw milk using Fourier transform infrared spectroscopy. *Journal of Food Science and Technology, 54*(8), 2394–2402. https://doi.org/10.1007/s13197-017-2680-y.

De La Fuente, M. A., & Juárez, M. (2005). Authenticity Assessment of Dairy Products. *Critical Reviews in Food Science and Nutrition, 45*(7–8), 563–585. https://doi.org/10.1080/10408690490478127.

Finete, V. d. L. M., Gouvêa, M. M., Marques, F. F. d. C., & Netto, A. D. P. (2013). Is it possible to screen for milk or whey protein adulteration with melamine, urea and ammonium sulphate, combining Kjeldahl and classical spectrophotometric methods? Food chemistry, 141(4), 3649-3655. doi:https://doi.org/10.1016/j.foodchem.2013.06.046.

García, V., Mollineda, R. A., & Sánchez, J. S. (2009). Index of balanced accuracy: A performance measure for skewed class distributions. In: Araujo H., Mendonça A.M., Pinho A.J., Torres M.I. (eds) Pattern Recognition and Image Analysis. IbPRIA 2009. Lecture Notes in Computer Science, vol 5524. Springer, Berlin, Heidelberg.

Gholami, R., & Fakhari, N. (2017). Chapter 27 - Support Vector Machine: Principles, Parameters, and Applications. In P. Samui, S. Sekhar, & V. E. Balas (Eds.), Handbook of Neural Computation (pp. 515-535): Academic Press.

Gurbanov, R., Gozen, A. G., & Severcan, F. (2018). Rapid classification of heavy metal-exposed freshwater bacteria by infrared spectroscopy coupled with chemometrics using supervised method. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 189*, 282–290. https://doi.org/10.1016/j.saa.2017.08.038.

Heck, J. M. L., van Valenberg, H. J. F., Dijkstra, J., & van Hooijdonk, A. C. M. (2009). Seasonal variation in the Dutch bovine raw milk composition. *Journal of Dairy Science, 92*(10), 4745–4755. https://doi.org/10.3168/jds.2009-2146.

International Organization for Standardization [ISO]. (2009). Milk - Determination of freezing point - Thermistor cryoscope method (ISO standard No. 5764). Retrived from https://www.iso.org/standard/43986.html.

International Organization for Standardization [ISO]. (2014). Milk and milk products — Determination of nitrogen content. In Part 1: Kjeldahl principle and crude protein calculation (ISO standard No. 8968). Retrived from https://www.iso.org/obp/ui#iso:std:iso:8968:-1:ed-2:v1:en.

International Organization for Standardization [ISO]. (2015). Milk, milk products, infant formula and adult nutritionals — Determination of fatty acids composition — Capillary gas chromatographic method (ISO standard No. 16958). Retrived from https://www.iso.org/obp/ui#iso:std:iso:16958:ed-1:v1:en.

Jawaid, S., Talpur, F. N., Sherazi, S. T. H., Nizamani, S. M., & Khaskheli, A. A. (2013). Rapid detection of melamine adulteration in dairy milk by SB-ATR–Fourier transform infrared spectroscopy. *Food chemistry, 141*(3), 3066–3071. https://doi.org/10.1016/j.foodchem.2013.05.106.

Jha, S. N., Jaiswal, P., Borah, A., Gautam, A. K., & Srivastava, N. (2014). Detection and Quantification of Urea in Milk Using Attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy. *Food and Bioprocess Technology, 8*(4), 926–933. https://doi.org/10.1007/s11947-014-1455-y.

Kačuráková, M., & Mathlouthi, M. (1996). FTIR and laser-Raman spectra of oligosaccharides in water: Characterization of the glycosidic bond. *Carbohydrate Research, 284*(2), 145–157. https://doi.org/10.1016/0008-6215(95)00412-2.

Kačuráková, M., & Wilson, R. H. (2001). Developments in mid-infrared FT-IR spectroscopy of selected carbohydrates. *Carbohydrate Polymers, 44*(4), 291–303. https://doi.org/10.1016/S0144-8617(00)00245-9.

Kartheek, M., Smith, A. A., Muthu, A. K., & Manavalan, R. (2011). Determination of adulterants in food: A review. *Journal of Chemical and Pharmaceutical Research, 3*(2), 629–636.

Kendall, H., Kuznesof, S., Dean, M., Chan, M.-Y., Clark, B., Home, R., ... Frewer, L. (2019). Chinese consumer's attitudes, perceptions and behavioural responses towards food fraud. *Food Control, 95*, 339–351. https://doi.org/10.1016/j.foodcont.2018.08.006.

Li, S., Sijtsema, S. J., Kornelis, M., Liu, Y., & Li, S. (2019). Consumer confidence in the safety of milk and infant milk formula in China. *Journal of Dairy Science, 102*(10), 8807–8818. https://doi.org/10.3168/jds.2019-16638.

Liu, J., Ren, J., Liu, Z.-M., & Guo, B.-H. (2015). A new comprehensive index for discriminating adulteration in bovine raw milk. *Food Chemistry, 172*(Supplement C), 251–256. https://doi.org/10.1016/j.foodchem.2014.09.063.

Luykx, D. M. A. M., Cordewener, J. H. G., Ferranti, P., Frankhuizen, R., Bremer, M. G. E. G., Hooijerink, H., & America, A. H. P. (2007). Identification of plant proteins in adulterated skimmed milk powder by high-performance liquid chromatography—mass spectrometry. *Journal of Chromatography A, 1164*(1), 189–197. https://doi.org/10.1016/j.chroma.2007.07.017.

Mecker, L. C., Tyner, K. M., Kauffman, J. F., Arzhantsev, S., Mans, D. J., & Gryniewicz-Ruzicka, C. M. (2012). Selective melamine detection in multiple sample matrices with a portable Raman instrument using surface enhanced Raman spectroscopy-active gold nanoparticles. *Analytica Chimica Acta, 733*, 48–55. https://doi.org/10.1016/j.aca.2012.05.001.

Moore, J. C., Spink, J., & Lipp, M. (2012). Development and Application of a Database of Food Ingredient Fraud and Economically Motivated Adulteration from 1980 to 2010. *Journal of Food Science, 77*(4), R118–R126. https://doi.org/10.1111/j.1750-3841.2012.02657.x.

Moyer, D. C., DeVries, J. W., & Spink, J. (2017). The economics of a food fraud incident – Case studies and examples including Melamine in Wheat Gluten. *Food Control, 71*, 358–364. https://doi.org/10.1016/j.foodcont.2016.07.015.

Nascimento, C. F., Santos, P. M., Pereira-Filho, E. R., & Rocha, F. R. P. (2017). Recent advances on determination of milk adulterants. *Food chemistry, 221*, 1232–1244. https://doi.org/10.1016/j.foodchem.2016.11.034.

Nicolaou, N., Xu, Y., & Goodacre, R. (2010). Fourier transform infrared spectroscopy and multivariate analysis for the detection and quantification of different milk species. *Journal of Dairy Science, 93*(12), 5651–5660. https://doi.org/10.3168/jds.2010-3619.

Qlip. Payment criteria analyses. https://www.qlip.nl/en/component/content/article?id=358:payment-criteria-analyses. Last accessed 20-12-2019.

Rattray, W., & Jelen, P. (1996). Protein standardization of milk and dairy products. *Trends in Food Science & Technology, 7*(7), 227–234. https://doi.org/10.1016/0924-2244(96)10024-8.

Royal Friesland Campina (2019). Milk price explained. Last accessed 20-12-2019 https://www.frieslandcampina.com/en/milk-price-explained/.

Santos, P. M., Pereira-Filho, E. R., & Rodriguez-Saona, L. E. (2013). Rapid detection and quantification of milk adulteration using infrared microspectroscopy and chemometrics analysis. *Food chemistry, 138*(1), 19–24. https://doi.org/10.1016/j.foodchem.2012.10.024.

Singh, P., & Gandhi, N. (2015). Milk preservatives and adulterants: Processing, regulatory and safety issues. *Food Rev Int, 31*. https://doi.org/10.1080/87559129.2014.994818.

Sneddon, N., Lopez-Villalobos, N., Hickson, R., & Shalloo, L. (2013). *Review of milk payment systems to identify the component value of lactose, 73*, 33–36.

Sokolova, M., Japkowicz, N., & Szpakowicz, S. (2006). Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation. In: Sattar A., Kang B. (eds) AI 2006: Advances in Artificial Intelligence. AI 2006. Lecture Notes in Computer Science, vol 4304. Springer, Berlin, Heidelberg.

Spink, J., Moyer, D. C., & Speier-Pero, C. (2016). Introducing the Food Fraud Initial Screening model (FFIS). *Food Control, 69*, 306–314. https://doi.org/10.1016/j.foodcont.2016.03.016.

Tripathy, S., Ghole, A. R., Deep, K., Vanjari, S. R. K., & Singh, S. G. (2017). A comprehensive approach for milk adulteration detection using inherent bio-physical properties as 'Universal Markers': Towards a miniaturized adulteration detection platform. *Food Chemistry, 217(Supplement, C)*, 756–765. https://doi.org/10.1016/j.foodchem.2016.09.037.

USDA-NASS. (2019). USDA National Agricultural Statistics Service. https://www.nass.usda.gov/. Last accessed 20-12-2019.

Yang, Y., Huisman, W., Hettinga, K. A., Liu, N., Heck, J., Schrijver, G. H., ... van Ruth, S. M. (2019). Fraud vulnerability in the Dutch milk supply chain: Assessments of farmers, processors and retailers. *Food Control, 95*, 308–317. https://doi.org/10.1016/j.foodcont.2018.08.019.

Zhang, J., Liang, L., Tian, Z., Chen, L., & Subirade, M. (2012). Preparation and in vitro evaluation of calcium-induced soy protein isolate nanoparticles and their formation mechanism study. *Food chemistry, 133*(2), 390–399. https://doi.org/10.1016/j.foodchem.2012.01.049.

Zuivelverordening (2000). Uitbetaling van boerderijmelk naar kwaliteit, samenstelling en gewicht. https://wetten.overheid.nl/BWBR0010719/2005-01-01/0/HoofdstukI, Last accessed 26-2-2020.