Food fraud vulnerability assessments and product authentication in Dutch and Chinese milk supply chains

Yuzheng Yang

Propositions

1. Good comprehension of food fraud vulnerability is essential to combat food fraud.
   (this thesis)
2. Developing authentication models separately for different milk types improves the performance of models.
   (this thesis)
3. Successful fraudsters are more innovative and creative than regulators.
4. Data collection, including samples acquisition and analytical measurements, is as important as data analysis in model development.
5. People tend to underestimate the influence of a good dress on working from home.
6. A pandemic leads to more rapid innovation in communication.

Propositions belonging to the thesis, entitled

Food fraud vulnerability assessments and product authentication in Dutch and Chinese milk supply chains

Yuzheng Yang

Wageningen, 16 April 2021

Food fraud vulnerability assessments and product authentication in Dutch and Chinese milk supply chains

Yuzheng Yang

Thesis committee

Promotors

Prof. Dr S.M. van Ruth
Special Professor, Food Authenticity and Integrity
Wageningen University & Research

Prof. Dr W. Huisman
Professor of Criminology
VU Amsterdam

Co-promotor

Dr K.A. Hettinga
Associate Professor, Food Quality and Design
Wageningen University & Research

Other members

Dr H.W.I. van Herpen, Wageningen University & Research
Prof. Dr N. Lord, University of Manchester, UK
Dr A. Haandrikman, Ausnutria B.V., Zwolle
Dr D. Wu, Queen's University Belfast, UK

Food fraud vulnerability assessments and product authentication in Dutch and Chinese milk supply chains

Yuzheng Yang

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 16 April 2021
at 11 a.m. in the Aula.

Table of Contents

CHAPTER 1
General introduction

1.1. Introduction

Milk is one of the most widely produced agricultural commodities worldwide. In the last three decades, the global milk production has increased from 530 million tons in 1988 to 843 million tons in 2018 (FAOSTAT, 2018). Asia is the leading milk producer (30%), followed by the European Union (28%), and north and central America (18%) (Sudhakaran & Minj, 2020). Especially east and southeast Asian countries contributed considerably to the increase in milk production in Asia. In China, the annual milk production reached 36 million tons in 2016, which was quadruplicated compared with the 8 million tons produced in 2000 (Li, 2016). With industrialization and urbanization, milk production and processing have both become more concentrated. For instance, the past practice of farmers trading milk within villages in China has changed towards selling raw milk to dairy processors. The structure of the milk industry organization differs among countries around the world. For instance in the Netherlands and Ireland, farmers' cooperatives own large-scale processing facilities, while in America and China farmers and processors mainly do business through individual contracts. Many countries consume most of their locally produced and processed dairy products themselves, while others, led by amongst others New Zealand, Germany, and the Netherlands, export a large proportion of their dairy production. Dairy chains connect the actors and activities involved in delivering milk products to the consumer, where each activity increases the value of the products.

In parallel with milk production, the demand for milk and milk products in many countries has also greatly increased, particularly in parts of east and southeast Asia (Handford, Campbell, & Elliott, 2016). The consumption of dairy products in China has increased by an average of 4.5% annually since 2011, reaching 36 kg milk equivalent per capita in 2016. Among the various dairy products, liquid milk products are the most consumed dairy products in east and southeast Asia (Li, 2016). Several factors contributed to the increased demand of milk products in developing countries, including economic growth, increase of public interest in high-protein diets, and government support of milk consumption through school milk programs (Handford et al., 2016). Despite the rapidly increased milk consumption in these countries, it is still much lower than that in Europe. In the Netherlands, for instance, the dairy consumption per capita was 130 kg milk equivalent in 2018, with drinking milk and cheese being the most consumed dairy products (Zuivel NL, 2018).

Milk comprises principally water in which various nutritional components, including fat, protein, lactose, minerals, and vitamins, are present. Milk is considered an oil-in-water emulsion, with the dispersed phase comprising milk fat and casein micelles, and the continuous phase being called milk serum (Huppertz & Kelly, 2006). The typical milk composition for lowland breeds, and the approximate variation in this composition, are presented in Table 1.1. Among the main milk components, fat has the largest variation in its content, followed by protein, whereas the lactose content is relatively stable. The variation in milk fat content results from several factors. Of the common breeds, milk from Jersey cows contains the highest fat content and from Holstein/Friesians cows the lowest (Fox, McSweeney, & Paul, 1998). In temperate climates, milk fat content is highest in winter and lowest in summer, mainly due to

difference in feeding between seasons. Normally, milk contains about 3.3 g/100 g protein, which comprises mainly casein and whey proteins. Lactose is the major carbohydrate in milk. Other carbohydrates, including oligosaccharides, glucose, and galactose, are present in trace amounts. Although the lactose concentration decreases progressively during lactation, the variation in lactose content of milk is much smaller than that of the other two macronutrients. This is mainly due to the important role of lactose in the osmotic pressure of milk. The osmotic pressure of milk is identical to that of blood, and is therefore very constant (Paques & Lindner, 2019). The minerals in milk include, amongst others, calcium, phosphorus, magnesium, potassium, selenium, and zinc. Although they are present at low concentrations, many minerals have important functions in the body, e.g. regarding bone formation and enzyme activities.

Table 1.1. General composition of milk from lowland breeds *.

Component	Average content (g/100 g)	Range (g/100 g)
Water	87.1	85.3-88.7
Fat	4.0	2.5-5.5
Protein	3.3	2.3-4.4
Lactose	4.6	3.8-5.3
Minerals	0.7	0.6-0.8
Organic acids	0.17	0.12-0.21
Miscellaneous	0.13	-

* Adapted from Walstra, Wouters, & Geurts (2005).

1.2. Food fraud

Food fraud incidents can be traced back to thousands of years ago. Laws in the *Tang* dynasty in ancient China regulated that offenders who sold adulterated foods or drugs would be imposed severe penalties (Zhangsun & Fang, 653). In Europe, there were for example laws in ancient Rome and Athens against the adulteration of wines with colours and flavours (Handford et al., 2016). Today, the food supply chain is getting more complex since both local and international trades are increasing. As a consequence, it is a big challenge to ensure the integrity of foods. One of the most widely accepted definitions of food fraud is as follows: food fraud is a collective term used to encompass the deliberate and intentional substitution, addition, tampering, or misrepresentation of food, food ingredients, or food packaging; or false or misleading statements made about a product, for economic gain (Spink & Moyer, 2011). The main difference with other food integrity issues is that food fraud is economically motivated (see Fig. 1.1). Food fraud could be practiced at multiple levels of the supply chain, i.e. from the primary producers to the end of the food chain, e.g. the catering service sector. Moreover, food fraud may involve simple procedures, for instance the dilution of a liquid food product, but may also include more complex adulteration operations. Food fraud can result in multiple consequences, such as financial and reputational losses for the food companies, a health threat to consumers, loss of public confidence for the whole industry, and even destruction of the image of an entire country. There has been an increasing awareness to combat food fraud. For instance, the Global Food Safety Initiative (GFSI) stressed the need to assess food fraud vulnerability and to develop fraud mitigation plans, in parallel with the food safety and food
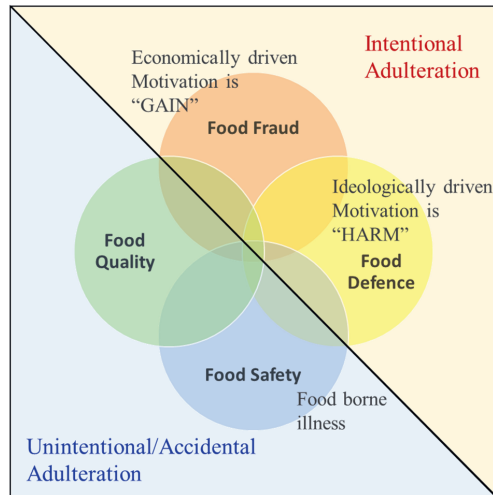
defence counter-actions (GFSI, 2014).

Fig. 1.1. Food risk types. Adapted from Global Food Safety Initiative (GFSI) position on mitigating the public health risk of food fraud (GFSI, 2014).

1.3. Food fraud risks and vulnerability assessments

1.3.1. Food fraud risk and vulnerability

Food fraud risk is considered the combination of likelihood and consequences of food fraud (Spink, Ortega, Chen, & Wu, 2017). A proper guidance for fraud prevention and control helps to not only reduce the likelihood of fraud, but also to mitigate its consequences. Understanding the intentional nature of food fraud leads to a focus shift from a safety-based approach to a fraud prevention and vulnerability reduction approach (van Ruth, Huisman, & Luning, 2017).

Vulnerability is defined as a weakness that creates an opportunity for an undesirable incident, and food fraud vulnerability is the susceptibility of a food system to fraud (Spink et al., 2017). In order to develop fraud prevention strategies, it is necessary to identify the potential factors which affect fraud vulnerability. Considering food fraud is resulting from criminal behaviour, a tested and validated criminological theory, the routine activities theory, has been applied to explain the occurrence of food fraud and theoretically frame food fraud vulnerability (van Ruth et al., 2017). The routine activities theory considers crime as an outcome of aggregation of suitable targets, motivated offenders, and the absence of capable guardians (Cohen & Felson, 1979). Based on this theory, food fraud vulnerability is defined by three corresponding elements: opportunities, motivations, and control measures (van Ruth et al., 2017). Coleman (1987) defined opportunity as a potential course of action, made possible by a particular set of social conditions, which has been symbolically incorporated into an actor's repertoire of behavioural possibilities, and defined motivation as to consist of a set of symbolic constructions defining certain kinds of goals and activities as appropriate and desirable and

others as lacking those qualities. In this study, food fraud opportunities refer to the circumstances that allow fraud to occur, and food fraud motivation to the fraudster's mindset toward committing fraud. The control measures are the actions or activities that may counteract the vulnerabilities caused by opportunities and motivations, eliminating or reducing opportunities and motivations. Following this theory, detailed factors underlying the three elements have been identified (van Ruth et al., 2017).

1.3.2. Detailed factors of fraud vulnerability

1.3.2.1. Fraud opportunities

Fraud opportunities concern technical opportunities, and opportunities in time and place. Technical opportunities related to the complexity of food composition and adulteration, and the difficulty in detecting or confirming such adulterations, while opportunities in time and place related to the processes of production and distribution, that may affect access to locations where fraud can be committed (van Ruth et al., 2017). From a technical perspective, the simplicity of adulteration and availability of knowledge to conduct adulteration will enable fraudsters to manipulate foods in practice. Therefore, they are considered as fraud factors for vulnerability. For instance, milk is a liquid product with a complex composition, and the physical state makes it easy to adulterate. In addition, if adulterated products cannot be detected by simple methods like visual recognition, then its susceptibility to fraud will increase. Consequently, detectability of adulteration is also a factor of fraud opportunities. Besides the technical opportunities, fraud vulnerabilities may be present by means of exposing opportunities in time and place to fraudsters to manipulate processes of production and distribution. For a complex dynamic supply chain, transparency and good relations within the chain are crucial for ensuring food integrity (Ali & Suleiman, 2018; Trienekens, Wognum, Beulens, & van der Vorst, 2012), which are thus considered as factors affecting fraud vulnerability. If a food product has been widely reported to be involved in fraud incidents, it is likely to exhibit opportunities for fraud as well.

1.3.2.2. Fraud motivations

Fraud motivations concern economic drivers as well as cultural and behavioural factors. Food fraud is for economic gains, therefore, the factors related to the price of product, for instance supply and price of materials, value-added attributes of the products, and competition level, may motivate offenders to gain extra profit under certain circumstances, and can affect fraud vulnerability. This was supported by the fact that before the melamine fraud incidents with infant formula in China in 2008, the demand of raw milk was exceeding the supply by far, leading to fierce competition for raw milk among dairy processors and soaring raw milk prices (Xiu & Klein, 2010). In addition, poor economic health of the organizations within a sector is likely to result in difficulties to achieve financial goals (Wang & Winton, 2012), which therefore can also motivate offenders to commit fraud.

As shown in research on crimes committed in occupational and business settings, potential offenders may be influenced by the organizational strategy and business culture.

Organizational strategy refers to the goals of the organization, and by what means organizational members should attained them. Business culture refers to behaviours expressing values and norms, expectation, attitudes, beliefs, and ideas that are shared by most of the organizational members (Huisman, 2016). When organizational strategies are short-termed and difficult to reach, they may be a motivation to commit an occupational crime. This way of thinking is suitable for food fraud, due to the fact that food fraud is also a criminal behaviour which occurs within organizations. Moreover, organizations that have been involved in fraud previously are more likely to be involved in fraud in the future again (Baucus & Near, 1991); consequently, previous corporate crimes may be an indicator of future reoccurrence. Fraud vulnerability can be impacted by the corruption level of the country, because a higher general corruption level in a country may be a proxy for firms to gain profit by illegal means (Martin, Cullen, Johnson, & Parboteeah, 2007). In addition, in countries with higher corruption levels, individual offenders may perceive fraud as a common and acceptable way to achieve their goals (Alibux, 2016). The corruption level, as a proxy for the prevalence of economic crimes, is a relevant fraud vulnerability factor.

1.3.2.3. Control measures

Control measures help to counteract the vulnerabilities resulting from opportunities and motivations, and are divided in technical and managerial controls. The technical controls refer to the actions related to testing or recording. For instance, a proper fraud monitoring system can enhance the detectability of adulteration by means of specifying fraud monitoring procedures, evidence-based sampling plans, anomaly detection methods, etc. Traceability systems allow food businesses to quickly target the affected products when something goes wrong. For example, a stock control system can ensure that the amount of raw material equates to the final product, which can help to discover flaws in mass balance and indicate losses or fraud occurrence. Another example of a technical control is a tracking and tracing system, which enables manufacturers not only to manage logistics efficiently, but also to respond rapidly in a product recall situation, e.g. in case of fraud occurrence.

Managerial controls are behaviour influential factors that may affect fraud vulnerability by reducing motivations for fraud of employees, suppliers, and other stakeholders. Managerial controls within organizations include ethical codes of conduct, whistle-blowing systems, integrity screening of employees, etc. It has been shown that a proper business code for ethical behaviour leads to less unethical behaviour (Kaptein, 2011); this can thus be considered as a factor for fraud vulnerability. A well-established whistle-blowing system, which contains standardized procedures and protects whistle-blowers, can increase the likelihood that employees disclose unethical behaviour, and is thereby crucial to mitigate food fraud (Soon & Manning, 2017). External managerial controls concern social control and transparency, industrial guidelines, and law enforcement, which are just as important as internal controls. Social control is an active or passive process of a group regulating itself according to the principles and values, and is important in fraud controls too. The group may refer to any actor within a food supply chain, or a specific tier such as organic farmers. Social control may occur

in forms of exchanging information, making decisions and fulfilling promises (van Ruth et al., 2017). Elliott (2014) pointed out that food fraud is hardly a priority for legislation and enforcement. In most EU countries, food authorities still need to refer potential criminal cases to the police, and only twelve of 28 member states have some units specifically dedicated to addressing food fraud (Gussow, 2020). Social control and industrial guidance, as well as legislation and enforcement on ensuring food integrity, play additional roles in fraud control.

1.3.3. *Food fraud vulnerability assessments*

In accordance with the three fraud elements and the detailed factors that may affect fraud vulnerability, Safe Supply of Affordable Food Everywhere (SSAFE) developed a food fraud vulnerability assessment (FFVA) tool, in collaboration with Wageningen University and VU University of Amsterdam. This tool consists of 50 questions, including eleven related to fraud opportunities, 20 to fraud motivations and 19 to control measures (SSAFE, 2017). The FFVA tool also takes into account the influence of the environment, which is in line with the suggestions of other scholars, that fraud mitigation strategy should operate at three levels, i.e. at organizational, supply chain, and global levels (Manning, 2016). The related fraud factors are assessed in both the internal and external environment of the business. In addition, it is stressed that the company's ability decreases along the dimension of the environmental layers (Fig. 1.2). The SSAFE FFVA tool is suited for self-assessment of individual business, independent of their size and geographic location. The tool can also be applied for evaluation and comparison of multiple tiers of food supply chains, to enable the profiling of fraud vulnerability for whole food supply chains. For instance, the FFVA tool was used for profiling and comparing fraud vulnerability for the main tiers of several food supply chains (Silvis, van Ruth, van der Fels-Klerx, & Luning, 2017; Yan, Erasmus, Aguilera Toro, Huang, & van Ruth, 2020; van Ruth, Luning, Silvis, Yang, & Huisman, 2018), comparing fraud vulnerability for different types of businesses in food service sector (van Ruth, van der Veeken, Dekker, Luning, & Huisman, 2020), and evaluating fraud vulnerability for specific food chains such as organic foods (van Ruth & de Pagter-de Witte, 2020).
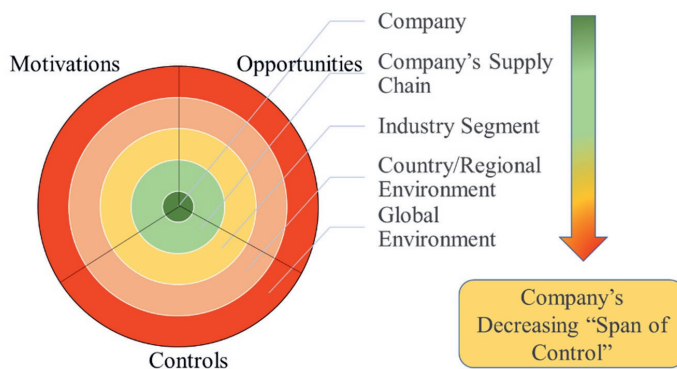

Fig. 1.2. The environment of the business and the three elements of food fraud. Adapted from the Introduction to the Safe Supply of Affordable Food Everywhere (SSAFE) food fraud vulnerability assessment tool (SSAFE, 2017).

1.4. Milk adulteration

Milk adulteration has a long history. For example, the New York Times (1858) reported that a swill milk scandal killed thousands of infants in New York in 1858. A hundred and fifty years later, the outbreak of melamine contamination of infant formula in China has again raised concerns globally. The infant formula adulterated with melamine caused the deaths of six infants in China, on top of which 300,000 children suffered kidney problems (Ministry of Health P. R. China, 2009). As a nitrogen-rich compound, melamine was added to mask the dilution of milk, which went unnoticed by the traditional analysis on protein content. The toxicity of melamine is linked to its inability to dissolve easily, which may result in acute renal failure (Pei et al., 2011). This melamine contamination incident was considered to be the result of the rapid and unregulated development of the Chinese dairy sector, and provoked a comprehensive reform of the Chinese food safety regime (Pei et al., 2011).

This infant formula incident raised public concerns regarding the integrity of milk and milk products. Based on several studies that analysed food fraud incidents in the past, milk was repeatedly reported as one of the most frequently adulterated foods (Moore, Spink, & Lipp, 2012; Zhang & Xue, 2016). Milk adulteration can, amongst others, be conducted by increasing the volume (dilution) or increasing the valuable components, where the fraud approach usually depends on the payment schemes' focus. Water dilution is the simplest approach of milk adulteration, which is probably the reason why, among various types of milk adulterations, water was found to be the most common adulterant (Barham, Khaskheli, Soomro, & Nizamani, 2014; Handford et al., 2016; Souza et al., 2011). Addition of other adulterants can be used to either mask the dilution with water or to increase the (apparent) level of components that are part of the payment scheme (e.g. protein, fat, lactose). Examples of adulterants focused on the valuable content include protein-rich substances such as milk powder and whey powder, nitrogen-rich compounds such as urea and melamine, and compounds that increase total solids content such as starch and sucrose (Pei et al., 2011; Schoder, 2010; Souza et al., 2011). Fat is another major milk component which has been frequently targeted by fraudsters (Nascimento, Santos, Pereira-Filho, & Rocha, 2017). The main fat adulterants include vegetable oils such as soybean, palm, coconut, and sunflower oil, and animal fats such as pork lard and beef tallow. Some detergents have been used for emulsification of external fat in the milk, and therefore are usually added together with the fat adulterants (Nascimento et al., 2017). In addition to the adulterants mentioned above, which aim to increase the apparent content of valuable milk components, some substances, that can extend the product shelf-life by decreasing microbial growth, have been used as milk adulterants (Singh & Gandhi, 2015). These adulterants include, amongst others, formalin, hydrogen peroxide, salicylic acid, bronopol, and sodium carbonate. These preservatives are either forbidden or its use is strictly limited in milk production. Most biological adulterants, like whey powder, starch, and vegetable oil, mainly cause a decrease in nutritional value, while other adulterants, including various types of chemicals, may also result in health problems of the consumers.

1.5. Detection of milk adulteration

Detection of milk adulteration can be achieved with a large range of analytical techniques. Generally, milk authentication concerns both confirmatory techniques, that aim to measure specific milk components or detect specific adulterants, and screening techniques, that aim to obtain a more comprehensive insight of samples by a fingerprinting approach, where no specific target needs to be selected beforehand.

1.5.1. Confirmatory techniques

Traditional milk adulteration detection strategies mainly aim at specific features of milk (Table 1.2). For instance, determination of freezing point and osmolality is used for detection of milk dilution (Büttel, Fuchs, & Holz, 2008; International Organization for Standardization [ISO], 2009), and the iodine test can be used for detection of milk adulterated with starch (Pradeep, Lakshminarayana, Varsha, & Kota, 2016). Considering that nitrogen-rich compounds are usually added to mask dilution by increasing the apparent protein content, which cannot be distinguished by traditional nitrogen-based protein analysis (Kjeldahl or Dumas methods), a number of techniques, including gas chromatography (GC), high performance liquid chromatography (HPLC), and mass spectroscopy (MS) have been applied for identification and quantification of nitrogen-rich adulterant compounds. For example, HPLC was used for detection of caseinomacropeptides to identify rennet whey addition (Ferreira & Oliveira, 2003), and GC-MS/MS was used for analysing the presence of melamine and its analogues in infant formula samples (Braekevelt, Lau, Feng, Ménard, & Tittlemier, 2011). Non-dairy fats can be detected by analysing, for instance, the triglyceride composition (ISO, 2019), fatty acids (Ntakatsane, Liu, & Zhou, 2013; Rebechi, Vélez, Vaira, & Perotti, 2016), and the sterol fraction (Alonso, Fontecha, Lozada, & Juárez, 1997), using, amongst others, GC or fluorescence spectroscopy. Identification of stable isotope ratios of atoms, such as carbon and hydrogen, using isotope ratio mass spectrometry (IR-MS) can specify the geographical origin of the products (Luo et al., 2016). Deoxyribonucleic acid (DNA)-based techniques are applied for distinguishing dairy products derived from specific animal species. For instance, real-time polymerase chain reaction (RT-PCR) was reported to be useful for the detection of camel milk that was adulterated with bovine milk (Wang et al., 2020). The above methods are some examples of techniques amongst many more analytical methods that can be used for confirmation of specific types of adulteration.

Table 1.2. Examples of confirmatory techniques for milk authentication [a].

Targeted parameters	Analytical instrument	Applications	Reference
Casein macropeptide	RP-HPLC	Detection of rennet whey in milk powder	(Ferreira & Oliveira, 2003)
C, H, O, N stable isotope ratios	IR-MS	Discrimination of geographical origins of milk	(Luo et al., 2016)
C, N stable isotope ratios	IR-MS	Discrimination organic milk from conventional milk	(Chung, Park, Yoon, Yang, & Kim, 2014)
DNA	RT-PCR	Determination of camel milk in bovine milk	(Wang et al., 2020)
Fatty acids	Fluorescence spectroscopy	Detection of vegetable oils in milk fat	(Ntakatsane et al., 2013)
Fatty acids	GC	Detection milk fat adulteration with animal fat	(Rebechi et al., 2016)
Freezing point	Thermistor cryoscope	Determination of extraneous water in milk	(ISO, 2009)
Maltodextrin	HPLC	Determination of maltodextrin in raw milk	(Moraes et al., 2017)
Melamine	ESI-MS/MS	Determination of melamine in milk powder	(Domingo et al., 2015)
Melamine	GC-MS	Identification and quantification of melamine in liquid milk and infant formula	(Lutter et al., 2011)
Melamine	HPLC	Determination of melamine in UHT milk	(Finete, Gouvêa, Marques, & Pereira Netto, 2015)
Melamine	MALDI-ToF-MS	Rapid detection of melamine in milk	(Su et al., 2013)
Melamine	RP-HPLC	Determination of melamine in dairy products	(Filazi, Sireli, Ekici, Can, & Karagoz, 2012)
Melamine and its analogues	GC-MS/MS	Determination of melamine and its analogy in infant formula	(Braekevelt et al., 2011; Miao et al., 2009)
Osmolality	Osmometer	Determination of additional water in milk	(Büttel et al., 2008)
Sterol fraction	GC	Detection of vegetable oils in milk fat	(Alonso et al., 1997)
Sugars	HPLC	Detection milk adulteration with soymilk	(Sharma, Rajput, Poonam, Dogra, & Tomar, 2009)
Triglyceride	GC	Detection of foreign fats (vegetable oils and animal fats) in milk	(ISO, 2019)
Urea	GC-IDMS	Quantification of urea in milk	(Dai et al., 2010)

[a] DNA, Deoxyribonucleic acid; ESI-MS/MS, electrospray ionisation-tandem mass spectrometry; GC-IDMS, gas chromatography-isotope dilution mass spectrometry; GC-MS, gas chromatography-mass spectrometry; IR-MS, isotope ratio mass spectrometry; MALDI-ToF-MS, matrix-assisted laser-desorption ionization-time of flight-mass spectrometry; RP-HPLC, reverse phase-high performance liquid chromatography; RT-PCR, real time-polymerase chain reaction.

1.5.2. Screening techniques

In recent years, with the increasing demand for detection of a broader range of adulterants, several rapid, non-destructive, and high throughput screening techniques, combined with chemometrics, have come into use for milk authentication. Vibrational spectroscopy is one of the most widely studied fingerprinting techniques for milk authentication. Some examples of fingerprinting methods are presented in Table 1.3. Milk adulteration with foreign proteins like whey and soy protein were successfully detected by near infrared spectroscopy (NIRS), Fourier transform infrared (FTIR) spectroscopy, and ultra-high performance liquid chromatography (UHPLC) (Jaiswal et al., 2015; Kasemsumran, Thanapase, & Kiatsoonthon, 2007; Xu, Yan, Cai, Wang, & Yu, 2013). Non-dairy fats like vegetable oils can be detected by direct analysis in real time ionization–high resolution mass spectrometry (DART-HRMS) (Hrbek, Vaclavik, Elich, & Hajslova, 2014). In addition, more complicated milk adulterations with multiple substances, including nitrogenous compounds and preservatives, were reported to be detectable by NIRS (Santos, Pereira-Filho, & Rodriguez-Saona, 2013), FTIR (Botelho, Reis, Oliveira, & Sena, 2015; Liu, Ren, Liu, & Guo, 2015), Raman spectroscopy (Nieuwoudt, Holroyd, McGoverin, Simpson, & Williams, 2016), and time domain nuclear magnetic resonance (TD-NMR) (Coimbra et al., 2020).

Despite the fact that most of the conventional screening analyses are laboratory-based, some analytical instruments have been miniaturized over the past decade to allow on-site tests. These techniques can offer user-friendly and rapid detection, and screen for the occurrence of suspect samples, prior to further laboratory confirmatory tests. Small-sized devices, including hand-held mid infrared (MIR) spectrometers and portable FTIR spectrometers (Santos et al., 2013), portable NIRS (Liu et al., 2018), and portable Raman spectrometers (Mecker et al., 2012), have been applied for screening of milk adulteration.

Besides the traditional screening methods, novel techniques which may complement existing methods have been explored for food authentication. An example of such a novel technique, that has not been used a lot yet for food authentication, is the application of sound. The advantage of this technique is that the measurement of acoustic properties usually requires little sample preparation, and is thus easy to use. Some examples of such techniques include broad acoustic resonance dissolution spectroscopy analysis (BARDS), and analysis of physical properties of sound propagation (van Ruth et al., 2019; Vos et al., 2016). Ultrasound measurements, including measurements of ultrasound velocity, attenuation, and other parameters, have been applied for characterization of dairy products (Dukhin, Goetz, & Travers, 2005), determination of the quality of reconstituted milk powder products (Meyer, Rajendram, & Povey, 2006), and microbial quality control of packaged milk (Gestrelius, Hertz, Nuamu, Person, & Lindström, 1993). Such techniques may thus also be applicable for rapid screening of milk adulteration.

Table 1.3. Examples of screening techniques for milk authentication [a].

Instrument	Statistics	Applications	Reference
ATR-FTIR	PCA, PLS, SIMCA	Determination of soymilk in liquid milk	(Jaiswal et al., 2015)
ATR-FTIR	PLS	Determination of melamine in liquid milk and milk powder	(Jawaid, Talpur, Sherazi, Nizamani, & Khaskheli, 2013)
DATR-HRMS	PCA, LDA	Detection of vegetable oils (rapeseed, sunflower, and soybean) in soft cheese	(Hrbek et al., 2014)
FTIR	PCA, SIMCA	Detection of milk adulteration with formaldehyde, hydrogen peroxide, bicarbonate, citrate, carbonate, starch, sucrose in milk	(Gondim, Junqueira, Souza, Ruisánchez, & Callao, 2017)
FTIR	PCA	Detection of milk adulteration with water, urea, melamine, ammonium chloride, whey, margarine, maltodextrin and milk powder in milk	(Liu et al., 2015)
FTIR	PLS-DA	Detection of milk adulteration with water, starch, sodium citrate, formaldehyde and sucrose	(Botelho et al., 2015)
NIR	PLS	Detection of foreign proteins (edible gelatine, industrial gelatine and soy protein) in yogurt	(Xu et al., 2013)
NIR	PLS, SIMCA	Detection of whey and water in milk	(Kasemsumran et al., 2007)
NIR	PLS-DA	Differentiation of organic milk from non-organic milk	(Liu et al., 2018)
NIR	SIMCA, PLS	Detection of milk adulteration with tap water, whey, urea, hydrogen peroxide and urea in milk	(Santos et al., 2013)
NMR	PCA, DA	Determination of geographic origin of cow milk	(Sacco et al., 2009)
Raman	PLS-DA	Detection of milk adulteration with small nitrogen-rich molecules and sucrose	(Nieuwoudt et al., 2016)
TD-NMR	PCA, PLS, SIMCA	Detection of formaldehyde in raw milk	(Coimbra et al., 2020)
UHPLC	PCA, SIMCA	Detection of adulteration of skimmed milk powder with foreign proteins	(Jablonski, Moore, & Harnly, 2014)

[a] ATR-FTIR, attenuated total reflectance-Fourier transform infrared spectroscopy; DART-HRMS, direct analysis in real time ionization-high resolution mass spectrometry; LDA, linear discriminant analysis; NIR, near infrared spectroscopy; PCA, principal component analysis; PLS, partial least squares; PLS-DA, partial least squares-discrimination analysis; SIMCA, soft independent modelling of class analogies; TD-NMR, time domain nuclear magnetic resonance; UHPLC, ultra-high performance liquid chromatography.

1.5.3. Data analysis

With the increasing complexity of the data produced by modern analytical instruments, especially those aimed at screening purposes, the statistical methods to be applied also become more diverse and complex. As a first step, descriptive analysis (e.g. calculation of means, standard deviation, correlations, etc) followed by inferential statistics (e.g. analysis of variances)

are probably the most frequently used methods (Granato, de Araújo Calado, & Jarvis, 2014; Nunes, Alvarenga, de Souza Sant'Ana, Santos, & Granato, 2015). Next, multivariate statistical analyses, also known as chemometrics, are often applied to analyse datasets containing numerous variables originating from the above described screening techniques (Brereton, 2015). Food authentication critically depends on the establishment of reference datasets that represent comprehensive and standardized information about the authentic foods. Chemometric tools are hence needed to handle the resulting large volume of data (Danezis, Tsagkaris, Camin, Brusic, & Georgiou, 2016). Some examples of unsupervised chemometric approaches are principal component analysis (PCA), and cluster analysis (CA) which are used for explorative analyses. Furthermore, various supervised techniques and associated classifiers have been applied, such as discriminant analysis (DA), partial least square DA, naive Bayes, support vector machines (SVM), soft independent modelling of class analogies (SIMCA), and k-nearest neighbour (kNN) approaches (Thinh, Thong, Cong, & Phong, 2019; Bouzembrak & Marvin, 2019; Tsakanikas, Karnavan, Panagou, & Nychas, 2020; Devos, Downey, & Duponchel, 2014; Aykas, Shotts, & Rodriguez-Saona, 2020). In addition, some deep learning algorithms such as convolutional neural networks, recurrent neural networks, and long short-term memory networks are also used for model development for food classification (Vo, Scanlan, & Turner, 2020; Hossain, Al-Hammadi, & Muhammad, 2018). The selection of the optimal data analysis approach is based on the specific research aim and dataset characteristics. It is important that the purpose and scope of the method should be explicitly defined before developing the model, because different purposes can influence the final acceptance criteria for the model performance (Alewijn, van der Voet, & van Ruth, 2016). For instance, the predictive performance of a model used for taking legal actions should be higher than that for on-site screening. Method validation allows an objective assessment of whether the model is fit for purpose, therefore it is an essential step of model development.

1.6. Knowledge gap and research justification

Recent food fraud incidents highlight the importance to combat fraud across food supply chains (Manning & Soon, 2014). Milk has been targeted by fraudsters for centuries with serious consequences. The fraud case of melamine in infant formula in China was one of the most significant food fraud incidents in modern history, causing domestic and international concern. Despite the fact that many analytical methods have been developed for milk authentication, it is unknown where, how, and how often milk supply chain networks are targeted. Understanding of the weaker and stronger actors in the milk supply chains, as well as the underlying factors, would help to comprehend the specific risks that are present in dairy chains in different countries, and would ultimately help to reduce fraud in the milk industry in a science-based and efficient manner.

The milk supply chains in the Netherlands and China are very different, and can be considered as the representatives of two typical structures of the dairy industry. On the one hand, the Dutch dairy industry was established hundreds of years ago and has developed over a long period, ending in a mature and cooperative-based dairy industry. On the other hand, the

Chinese dairy industry only started growing in recent decades, and is still in its infancy, but is undergoing rapid developments. Moreover, there were rarely incidents reported in the Netherlands over the past decades, while several incidents in the Chinese milk chain with severe consequences have been reported. It is, therefore, of great interest to get better insights into the fraud vulnerability in these two milk supply chains.

Considering the knowledge gaps mentioned above, this thesis focuses on liquid milk adulteration and deals with (a) vulnerabilities of milk supply chains in the Netherlands and China, and (b) the link between underlying vulnerability factors and fraud prevalence.

1.7. Research aim and thesis outline

The main objective of this thesis is to comprehend fraud vulnerabilities in two types of milk supply chains, represented by the Dutch and Chinese supply chains, their underlying factors, and the relationship of these factors with fraud prevalence. The sub-objectives are:

➢ To examine and critically compare fraud vulnerability profiles of businesses along the milk supply chains in the Netherlands and China;

➢ To explore the use of routine compositional analysis and a novel ultrasound-based method for detection of milk adulteration;

➢ To investigate the prevalence of milk adulteration in the Chinese market;

➢ To explore the relationship between fraud vulnerability profiles and prevalence of milk adulteration in China.

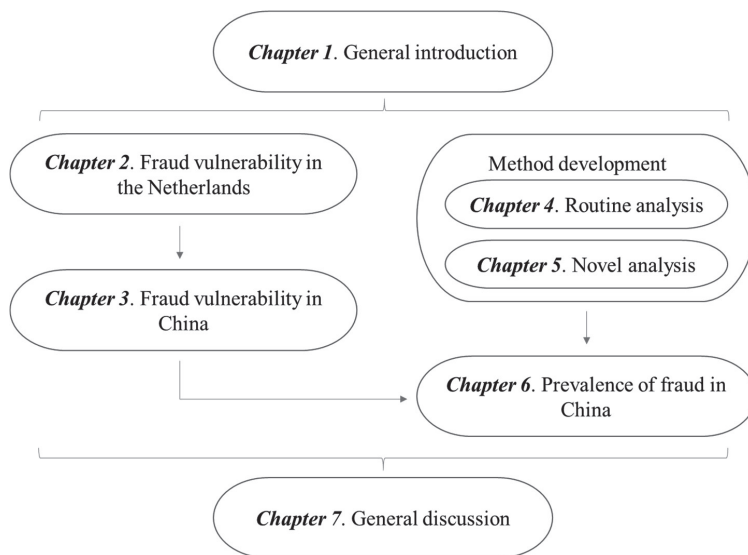The outline of thesis is graphically depicted in Fig. 1.3.

Fig. 1.3. Schematic overview of the thesis outline.

Chapter 1 provides a general introduction on the topics related to the thesis. Chapters 2 and 3 describe the evaluation of fraud vulnerabilities in milk supply chains in the Netherlands and China. In Chapter 2, the fraud vulnerability for the main tiers in the Dutch milk supply chain, namely dairy farmers, processors, and retailers, is assessed, and the similarities and differences in fraud vulnerability between the tier groups are identified. Furthermore, the vulnerabilities for the farmers applying different farming management systems, i.e. organic farms, intermediate pasture farms, and conventional farms are compared. In Chapter 3 the fraud vulnerability of dairy farmers and processors in the main milk production areas in China is examined. The differences in fraud vulnerability between the Chinese farmers and processors, and the differences between the Chinese farms of different scales and regions are presented.

Chapter 4 and Chapter 5 focus on development of analytical methods for milk fraud detection. Chapter 4 assesses the impact of common milk adulterants on the results of routine milk composition analysis by FTIR. The resulting data is analysed with both univariate and multivariate statistical approaches, so that the performance of routine compositional analysis in detection of common milk adulteration can be evaluated. In addition, the profitability of common milk adulteration is investigated, and profitable adulterations, which go unnoticed by the routine compositional analysis, are identified. Chapter 5 explores a novel, rapid, and non-destructive technique, based on a pulse-echo ultrasonic system, for characterisation of milk and cream samples and dilutions thereof. The ultrasound velocity, as well as the viscosity, density, and compositional properties are determined for six dairy products with different fat contents, and their water diluted counterparts. The ability of this novel approach to characterise liquid dairy products is evaluated, and the underlying causes for velocity differences between various dairy products are discussed.

Chapter 6 investigates the authenticity of milk in the Chinese market. Milk samples from different areas in China are tested by the authentication method developed in Chapter 4, and the prevalence of adulteration-suspected milk samples from several Chinese regions is determined. The prevalence of milk adulteration is further related to the fraud vulnerability profiles determined in Chapter 3, and the factors contributing to higher fraud prevalence in certain regions are identified.

The general discussion (Chapter 7) integrates the findings of this thesis. The fraud vulnerabilities in Dutch and Chinese milk supply chains are compared. Furthermore, the performance of the screening milk authentication method that have been presented in Chapters 4-6 are also compared for milks produced in the Netherlands and China, as well as the performance of the routine compositional analysis method and the novel ultrasound-based method. Furthermore, additional relations between fraud vulnerabilities and fraud prevalence are explored in this chapter. Finally, intervention strategies to monitor milk adulteration and mitigate fraud risks using the results of this thesis are proposed.

References

Alewijn, M., van der Voet, H., & van Ruth, S. (2016). Validation of multivariate classification methods using analytical fingerprints-concept and case study on organic feed for laying hens. *Journal of Food Composition and Analysis, 51*, 15-23.

Ali, M. H., & Suleiman, N. (2018). Eleven shades of food integrity: A halal supply chain perspective. *Trends in Food Science & Technology, 71*, 216-224.

Alibux, A. N. (2016). Criminogenic conditions, bribery and the economic crisis in the EU: A macro level analysis. *European Journal of Criminology, 13*(1), 29-49.

Alonso, L., Fontecha, J., Lozada, L., & Juárez, M. (1997). Determination of mixtures in vegetable oils and milk fat by analysis of sterol fraction by gas chromatography. *Journal of the American Oil Chemists' Society, 74*(2), 131-135.

Aykas, D. P., Shotts, M. L., & Rodriguez-Saona, L. E. (2020). Authentication of commercial honeys based on Raman fingerprinting and pattern recognition analysis. *Food Control, 117*, 107346.

Barham, G. S., Khaskheli, M., Soomro, A. H., & Nizamani, Z. A. (2014). Detection and extent of extraneous water and adulteration in milk consumed at Hyderabad, Pakistan. *Journal of Food and Nutrition Sciences, 2*(2), 47-52.

Baucus, M. S., & Near, J. P. (1991). Can illegal corporate behavior be predicted? An event history analysis. *Academy of Management Journal, 34*(1), 9-36.

Botelho, B. G., Reis, N., Oliveira, L. S., & Sena, M. M. (2015). Development and analytical validation of a screening method for simultaneous detection of five adulterants in raw milk using mid-infrared spectroscopy and PLS-DA. *Food Chemistry, 181*, 31-37.

Bouzembrak, Y., & Marvin, H. J. (2019). Impact of drivers of change, including climatic factors, on the occurrence of chemical food safety hazards in fruits and vegetables: A Bayesian network approach. *Food Control, 97*, 67-76.

Braekevelt, E., Lau, B. P. Y., Feng, S., Ménard, C., & Tittlemier, S. A. (2011). Determination of melamine, ammeline, ammelide and cyanuric acid in infant formula purchased in Canada by liquid chromatography-tandem mass spectrometry. *Food Additives & Contaminants: Part A, 28*(6), 698-704.

Brereton, R. G. (2015). Pattern recognition in chemometrics. *Chemometrics and Intelligent Laboratory Systems, 149*, 90-96.

Büttel, B., Fuchs, M., & Holz, B. (2008). Freezing point osmometry of milk to determine the additional water content – an issue in general quality control and German food regulation. *Chemistry Central Journal, 2*(1), 6.

Capuano, E., Elgersma, A., Tres, A., & van Ruth, S. M. (2014). Phytanic and pristanic acid content in Dutch farm milk and implications for the verification of the farming management system. *International Dairy Journal, 35*(1), 21-24.

Chung, I., Park, I., Yoon, J., Yang, Y., & Kim, S. (2014). Determination of organic milk authenticity using carbon and nitrogen natural isotopes. *Food Chemistry, 160*, 214-218.

Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American Sociological Review, 44*(4), 588-608.

Coimbra, P. T., Bathazar, C. F., Guimarães, J. T., Coutinho, N. M., Pimentel, T. C., Neto, R. P. C., & Cruz, A. G. (2020). Detection of formaldehyde in raw milk by time domain nuclear magnetic resonance and chemometrics. *Food Control, 110*, 107006.

Coleman, J. W. (1987). Toward an integrated theory of white-collar crime. *American Journal of Sociology, 93*(2), 406-439.

Dai, X., Fang, X., Su, F., Yang, M., Li, H., Zhou, J., & Xu, R. (2010). Accurate analysis of urea in milk and milk powder by isotope dilution gas chromatography-mass spectrometry. *Journal of Chromatography B, 878*(19), 1634-1638.

Danezis, G. P., Tsagkaris, A. S., Camin, F., Brusic, V., & Georgiou, C. A. (2016). Food authentication: Techniques, trends & emerging approaches. *Trends in Analytical Chemistry, 85*, 123-132.

Devos, O., Downey, G., & Duponchel, L. (2014). Simultaneous data pre-processing and SVM classification model selection based on a parallel genetic algorithm applied to spectroscopic data of olive oils. *Food Chemistry, 148,* 124-130.

Domingo, E. d. C., Tireli, A. A., Nunes, C. A., Batista, A. V., Guerreiro, M. C., & Pinto, S. M. (2015). Rapid extraction of melamine in powdered milk for direct electrospray ionization tandem mass spectrometry analysis. *Talanta, 132*, 535-540.

Dukhin, A. S., Goetz, P. J., & Travers, B. (2005). Use of ultrasound for characterizing dairy products. *Journal of Dairy Science, 88*(4), 1320-1334.

Elliott, C. (2014). *Elliott Review into the Integrity and Assurance of Food Supply Networks*. London: Department for Environment, Food & Rural Affairs.

FAOSTAT. (2018). FAOSTAT. Food and Agriculture Organization of the United Nations. Retrieved from http://www.fao.org/faostat/en/#data/QV.

Ferreira, I. M. P. L. V. O., & Oliveira, M. B. P. P. (2003). Determination of caseinomacropeptide by an RP-HPLC method and monitoring of the addition of rennet whey to powdered milk. *Journal of Liquid Chromatography & Related Technologies, 26*(1), 99-107.

Filazi, A., Sireli, U. T., Ekici, H., Can, H. Y., & Karagoz, A. (2012). Determination of melamine in milk and dairy products by high performance liquid chromatography. *Journal of Dairy Science, 95*(2), 602-608.

1

Finete, V. D. L. M., Gouvêa, M. M., Marques, F. F. D. C., & Pereira Netto, A. D. (2015). Validation of a method of high performance liquid chromatography with fluorescence detection for melamine determination in UHT whole bovine milk. *Food Control, 51*, 402-407.

Fox, P. F., McSweeney, P. L., & Paul, L. (1998). *Dairy Chemistry and Biochemistry*. London: Blackie Academic & Professional.

Gestrelius, H., Hertz, T. G., Nuamu, M., Person, H. W., & Lindström, K. (1993). A non-destructive ultrasound method for microbial quality control of aseptically packaged milk. *LWT - Food Science and Technology, 26*(4), 334-339.

Global Food Safety Initiative. (2014). GFSI position on mitigating the public health risk of food fraud*. In *Global Food Safety Initiative, Consumer Goods Forum*. Retrieved from https://mygfsi.com/wp-content/uploads/2019/09/Food-Fraud-GFSI-Position-Paper.pdf.

Gondim, C. d. S., Junqueira, R. G., Souza, S. V. C. d., Ruisánchez, I., & Callao, M. P. (2017). Detection of several common adulterants in raw milk by MID-infrared spectroscopy and one-class and multi-class multivariate strategies. *Food Chemistry, 230*(Supplement C), 68-75.

Granato, D., de Araújo Calado, V. M., & Jarvis, B. (2014). Observations on the use of statistical methods in food science and technology. *Food Research International, 55*, 137-149.

Gussow, K. E. (2020). *Finding Food Fraud: Explaining the Detection of Food Fraud in the Netherlands*. Doctoral dissertation, Free University Amsterdam, the Netherlands.

Handford, C. E., Campbell, K., & Elliott, C. T. (2016). Impacts of milk fraud on food safety and nutrition with special emphasis on developing countries. *Comprehensive Reviews in Food Science and Food Safety, 15*(1), 130-142.

Hossain, M. S., Al-Hammadi, M., & Muhammad, G. (2018). Automatic fruit classification using deep learning for industrial applications. *IEEE Transactions on Industrial Informatics*, *15*(2), 1027-1034.

Hrbek, V., Vaclavik, L., Elich, O., & Hajslova, J. (2014). Authentication of milk and milk-based foods by direct analysis in real time ionization-high resolution mass spectrometry (DART-HRMS) technique: A critical assessment. *Food Control, 36*(1), 138-145.

Huisman, W. (2016). Criminogenic organizational properties and dynamics. In *The Oxford Handbook of White-collar Crime*. Oxford: Oxford University Press.

Huppertz, T., & Kelly, A. L. (2006). Physical chemistry of milk fat globules. In Fox P. F. & McSweeney P. L. H. (Eds.), *Advanced Dairy Chemistry 2 Lipids*. Boston, MA: Springer.

International Organization for Standardization [ISO]. (2009). ISO 5764:2009 [IDF 108:2009] Milk — Determination of freezing point — Thermistor cryoscope method (Reference method). Retrieved from https://www.iso.org/standard/43986.html.
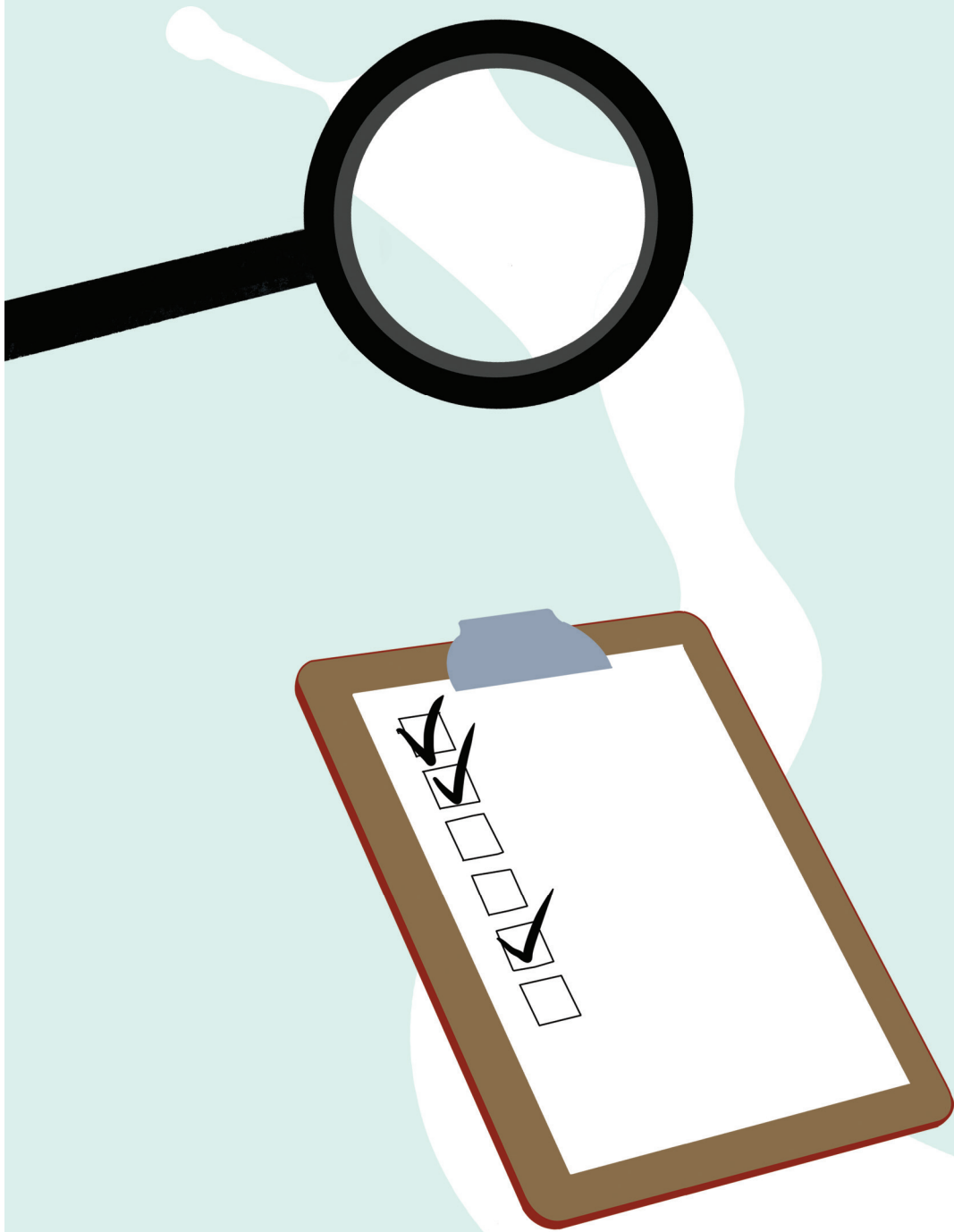
International Organization for Standardization [ISO]. (2019). ISO 17678:2019 [IDF 202:2019] Milk and milk products — Determination of milk fat purity by gas chromatographic analysis of triglycerides.Retrieved from https://www.iso.org/standard/72284.html.

Jablonski, J. E., Moore, J. C., & Harnly, J. M. (2014). Nontargeted detection of adulteration of skim milk powder with foreign proteins using UHPLC-UV. *Journal of Agricultural and Food Chemistry, 62*(22), 5198-5206.

Jaiswal, P., Jha, S. N., Borah, A., Gautam, A., Grewal, M. K., & Jindal, G. (2015). Detection and quantification of soymilk in cow-buffalo milk using attenuated total reflectance Fourier transform infrared spectroscopy (ATR-FTIR). *Food Chemistry, 168*, 41-47.

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.

Kaptein, M. (2011). Toward effective codes: Testing the relationship with unethical behavior. *Journal of Business Ethics, 99*(2), 233-251.

Kasemsumran, S., Thanapase, W., & Kiatsoonthon, A. (2007). Feasibility of near-infrared spectroscopy to detect and to quantify adulterants in cow milk. *Analytical Sciences, 23*(7), 907-910.

Li, S. (2016). White paper on China dairy 2016. Retrieved from http://www.sdddc.org/en/download/detail-251.aspx.

Liu, J., Ren, J., Liu, Z., & Guo, B. (2015). A new comprehensive index for discriminating adulteration in bovine raw milk. *Food Chemistry, 172*(Supplement C), 251-256.

Liu, N., Parra, H. A., Pustjens, A., Hettinga, K., Mongondry, P., & van Ruth, S. M. (2018). Evaluation of portable near-infrared spectroscopy for organic milk authentication. *Talanta, 184,* 128-135.

Luo, D., Dong, H., Luo, H., Xian, Y., Guo, X., & Wu, Y. (2016). Multi-element (C, N, H, O) stable isotope ratio analysis for determining the geographical origin of pure milk from different regions. *Food Analytical Methods, 9*(2), 437-442.

Lutter, P., Savoy-Perroud, M., Campos-Gimenez, E., Meyer, L., Goldmann, T., Bertholet, M., & Delatour, T. (2011). Screening and confirmatory methods for the determination of melamine in cow's milk and milk-based powdered infant formula: Validation and proficiency-tests of ELISA, HPLC-UV, GC-MS and LC-MS/MS. *Food Control, 22*(6), 903-913.

Manning, L. (2016). Food fraud: Policy and food chain. *Current Opinion in Food Science, 10*, 16-21.

Manning, L., & Soon, J. M. (2014). Developing systems to control food adulteration. *Food Policy, 49*, 23-32.

1

Martin, K. D., Cullen, J. B., Johnson, J. L., & Parboteeah, K. P. (2007). Deciding to bribe: A cross-level analysis of firm and home country influences on bribery activity. *Academy of Management Journal, 50*(6), 1401-1422.

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.

Meyer, S., Rajendram, V. S., & Povey, M. J. W. (2006). Characterization of reconstituted milk powder by ultrasound spectroscopy. *Journal of Food Quality, 29*(4), 405-418.

Miao, H., Fan, S., Wu, Y., Zhang, L., Zhou, P., Chen, H., Zhao, Y., & Li, J. (2009). Simultaneous determination of melamine, ammelide, ammeline, and cyanuric acid in milk and milk products by gas chromatography-tandem mass spectrometry. *Biomedical and Environmental Sciences, 22*(2), 87-94.

Ministry of Health P. R. China (2009). Routine press conference. Retrieved from https://archive.vn/20200714090708/http://www.scio.gov.cn/xwfbh/gbwxwfbh/fbh/Document/311421/311421.htm.

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.

Moraes, F. S., da Costa, M. P., de Melo Silva, V. L., de Barros Pinto Moreira, R. V., de Barros, R. F., Mársico, E. T., Conte-Junior, C. A., & de Oliveira Silva, A. C. (2017). Development of HPLC-ELSD method for determination of maltodextrin in raw milk. *Food Chemistry, 217*(Supplement C), 346-351.

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.

New York Times (1858). How we poison our children. Retrieved from https://www.nytimes.com/1858/05/13/archives/how-we-poison-our-children.html.

Nieuwoudt, M. K., Holroyd, S. E., McGoverin, C. M., Simpson, M. C., & Williams, D. E. (2016). Raman spectroscopy as an effective screening method for detecting adulteration of milk with small nitrogen-rich molecules and sucrose. *Journal of Dairy Science, 99*(4), 2520-2536.

Ntakatsane, M. P., Liu, X. M., & Zhou, P. (2013). Short communication: Rapid detection of milk fat adulteration with vegetable oil by fluorescence spectroscopy. *Journal of Dairy Science, 96*(4), 2130-2136.

Nunes, C. A., Alvarenga, V. O., de Souza Sant'Ana, A., Santos, J. S., & Granato, D. (2015). The use of statistical software in food science and technology: Advantages, limitations and misuses. *Food Research International, 75*, 270-280.

Paques, M., & Lindner, C. (2019). *Lactose: Evolutionary Role, Health Effects, and Applications*. Waltham, MA: Academic Press.

Pei, X., Tandon, A., Alldrick, A., Giorgi, L., Huang, W., & Yang, R. (2011). The China melamine milk scandal and its implications for food safety regulation. *Food Policy, 36*(3), 412-420.

Pradeep, S., Lakshminarayana, P., Varsha, R., & Kota, S. K. (2016). Screening of adulterants in milk. *International Journal of Current Research and Review, 8*(12), 25.

Rebechi, S. R., Vélez, M. A., Vaira, S., & Perotti, M. C. (2016). Adulteration of Argentinean milk fats with animal fats: Detection by fatty acids analysis and multivariate regression techniques. *Food Chemistry, 192*, 1025-1032.

Sacco, D., Brescia, M. A., Sgaramella, A., Casiello, G., Buccolieri, A., Ogrinc, N., & Sacco, A. (2009). Discrimination between southern Italy and foreign milk samples using spectroscopic and analytical data. *Food Chemistry, 114*(4), 1559-1563.

Santos, P. M., Pereira-Filho, E. R., & Rodriguez-Saona, L. E. (2013). Application of hand-held and portable infrared spectrometers in bovine milk analysis. *Journal of Agricultural and Food Chemistry, 61*(6), 1205-1211.

Schoder, D. (2010). Melamine milk powder and infant formula sold in east Africa. *Journal of Food Protection, 73*(9), 1709-1714.

Sharma, R., Rajput, Y. S., Poonam, Dogra, G., & Tomar, S. K. (2009). Estimation of sugars in milk by HPLC and its application in detection of adulteration of milk with soymilk. *International Journal of Dairy Technology, 62*(4), 514-519.

Silvis, I. C. J., van Ruth, S. M., van der Fels-Klerx, H. J., & Luning, P. A. (2017). Assessment of food fraud vulnerability in the spices chain: An explorative study. *Food Control, 81*, 80-87.

Singh, P., & Gandhi, N. (2015). Milk Preservatives and adulterants: Processing, regulatory and safety issues. *Food Reviews International, 31*(3), 236-261.

Soon Jan, M., & Manning, L. (2017). Whistleblowing as a countermeasure strategy against food crime. *British Food Journal, 119*(12), 2630-2652.

Souza, S. S., Cruz, A. G., Walter, E. H., Faria, J. A., Celeghini, R. M., Ferreira, M. M., & Sant'Ana, A. S. (2011). Monitoring the authenticity of Brazilian UHT milk: A chemometric approach. *Food Chemistry, 124*(2), 692-695.

Spink, J., & Moyer, D. C. (2011). Defining the public health threat of food fraud. *Journal of Food Science, 76*(9), R157-R163.

Spink, J., Ortega, D., Chen, C., & Wu, F. (2017). Food fraud prevention shifts the food risk focus to vulnerability. *Trends in Food Science & Technology, 62*, 215-220.

SSAFE. (2017). Food fraud vulnerability assessment tool. Retrieved from http://www.ssafe-food.org/our-projects/?proj=365.

Su, X., Zhou, H., Chen, F., Gao, B., Liu, Z., Zhang, Y., & Gao, Z. (2013). Modified SBA-15 matrices for high-throughput screening of melamine in milk samples by MALDI-TOF MS. *International Journal of Mass Spectrometry, 338*, 39-44.

Sudhakaran V, A., & Minj, J. (2020). Basic Facts About Dairy Processing and Technologies. In Minj J., Sudhakaran A., & Kumari A. (Eds.) *Dairy Processing: Advanced Research to Applications*. Singapore: Springer.

Thinh, N. T., Thong, N. D., Cong, H. T., & Phong, N. T. T. (2019). Mango classification system based on machine vision and artificial intelligence. In *2019 7th International Conference on Control, Mechatronics and Automation (ICCMA)*, Delft, Netherlands.

Trienekens, J. H., Wognum, P. M., Beulens, A. J. M., & van der Vorst, J. G. A. J. (2012). Transparency in complex dynamic food supply chains. *Advanced Engineering Informatics, 26*(1), 55-65.

Tsakanikas, P., Karnavas, A., Panagou, E. Z., & Nychas, G. J. (2020). A machine learning workflow for raw food spectroscopic classification in a future industry. *Scientific Reports, 10*(1), 1-11.

van Ruth, S. M., Hettinga, F., Dekker, P., & Fitzpatrick, D. (2019). The sound of the sand from the Dutch shores. *Applied Acoustics, 154*, 1-10.

van Ruth, S. M., Huisman, W., & Luning, P. A. (2017). Food fraud vulnerability and its key factors. *Trends in Food Science & Technology, 67*, 70-75.

van Ruth, S. M., Luning, P. A., Silvis, I. C. J., Yang, Y., & Huisman, W. (2018). Differences in fraud vulnerability in various food supply chains and their tiers. *Food Control, 84*(Supplement C), 375-381.

van Ruth, S. M., van der Veeken, J., Dekker, P., Luning, P. A., & Huisman, W. (2020). Feeding fiction: Fraud vulnerability in the food service industry. *Food Research International, 133*, 109158.

van Ruth, S. M., & de Pagter-de Witte, L. (2020). Integrity of organic foods and their suppliers: Fraud vulnerability across chains. *Foods, 9*(2)*,* 188.

Vo, S. A., Scanlan, J., & Turner, P. (2020). An application of convolutional neural network to lobster grading in the southern rock lobster supply chain. *Food Control*, *113*, 107184.

Vos, B., Crowley, S. V., O'Sullivan, J., Evans-Hurson, R., McSweeney, S., Krüse, J., Ahmed, R., Firzpatrick, D., & O'Mahony, J. A. (2016). New insights into the mechanism of rehydration of milk protein concentrate powders determined by broadband acoustic resonance dissolution spectroscopy (BARDS). *Food Hydrocolloids, 61*, 933-945.

Walstra, P., Wouters, J. T., & Geurts, T. J. (2005). *Dairy Science and Technology*. Boca Raton, FL: CRC press.

Wang, T. Y., & Winton, A. (2012). Competition and corporate fraud waves. Paper presented at *the 7th Annual Conference on Empirical Legal Studies*. November 9-10, 2012, Standford, CA. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id= 1783752.

Wang, Z., Li, T., Yu, W., Qiao, L., Liu, R., Li, S., Zhao, Y., Yang, S., & Chen, A. (2020). Determination of content of camel milk in adulterated milk samples by normalized real-time polymerase chain reaction system based on single-copy nuclear genes. *Journal of the Science of Food and Agriculture, 100*(8), 3465-3470.

Xiu, C., & Klein, K. K. (2010). Melamine in milk products in China: Examining the factors that led to deliberate use of the contaminant. *Food Policy, 35*(5), 463-470.

Xu, L., Yan, S., Cai, C., Wang, Z., & Yu, X. (2013). The feasibility of using near-infrared spectroscopy and chemometrics for untargeted detection of protein adulteration in yogurt: Removing unwanted variations in pure yogurt. *Journal of Analytical Methods in Chemistry*, *2013*, 201873.

Yan, J., Erasmus, S. W., Aguilera Toro, M., Huang, H., & van Ruth, S. M. (2020). Food fraud: Assessing fraud vulnerability in the extra virgin olive oil supply chain. *Food Control, 111*, 107081.

Zhang, W., & Xue, J. (2016). Economically motivated food fraud and adulteration in China: An analysis based on 1553 media reports. *Food Control, 67*, 192-198.

Zhangsun, W., & Fang, X. (653). *Tang Lyv Shu Yi* (In ancient Chinese). Retrieved from https://zh.m.wikisource.org/zh-hans/%E5%94%90%E5%BE%8B%E7%96%8F%E8%AD%B0.

Zuivel NL (2018). Dutch dairy in figures 2018. Retrieved from https://www.zuivelnl.org/uploads/images/ZIC2018-ENG.pdf.

1

CHAPTER 2

Fraud vulnerability in the Dutch milk supply chain: Assessments of farmers, processors and retailers

Yang, Y., Huisman, W., Hettinga, K. A., Liu, N., Heck, J., Schrijver, G. H., Gaiardoni, L., & van Ruth, S. M. (2019). Fraud vulnerability in the Dutch milk supply chain: Assessments of farmers, processors and retailers. *Food Control, 95*, 308-317.

Abstract

Food fraud surfaces regularly, anywhere in the world. Not only the companies involved in food fraud suffer from losses when food fraud occurs, other actors in the supply chain and branch of industry are often painted with the same brush. Milk has been a common fraud target in the past and, therefore, fraud is a concern for companies involved in milk production. In order to manage and prevent fraud in the milk supply chain, a good insight into the vulnerabilities of companies and their supply chain networks is pivotal. The aim of the current study is to understand (a) the fraud vulnerability of the general milk supply chain in the Netherlands and its tiers (farmers, processors, retailers) and (b) the differences in fraud vulnerability of farmers producing organic milk, green intermediate 'pasture milk' and conventional milk. The SSAFE food fraud assessment tool was slightly adapted to the milk supply chain and used to examine the fraud vulnerability of the 38 businesses of the three tiers in the study: 30 farmers, 4 milk processors and 4 retailers. Forty-eight fraud factors related to opportunities, motivations and control measures were examined. Subsequently, key fraud factors were identified. The three tier groups showed major similarities in motivation related fraud factors, and large differences in fraud opportunities and controls. There were also differences observed between the organic and non-organic farmers, with organic farmers being slightly more vulnerable than their non-organic counterparts. From this study it appears that the milk supply chain in the Netherlands is low to medium vulnerable to fraud but the key factors contributing to the vulnerability differ between the tiers (farmers, processors, retailers). Management of the fraud risks requires consideration of these differences.

Keywords

Dairy supply chain; Fraud factor; Fraud mitigation; Milk adulteration; Organic farm; Vulnerability assessment

2.1. Introduction

Food fraud has occurred throughout history with serious consequences. Food fraud is a collective term that includes deliberate and intentional substitution, addition, tampering or misrepresentation of food, food ingredients or food packaging; or false or misleading statements made about a product for economic gain (Spink & Moyer, 2011). The consequences of food fraud vary. It can cause both financial and reputational damages to directly affected companies, but also to all other actors in the supply chain and sector. Other stakeholders, such as national authorities and industrial associations, lose also public trust when a serious incident happens. Furthermore, food fraud can even do harm to human health under certain circumstances. The melamine scandal (2008) and horse meat scandal (2011) attracted considerable attention of the public. Since modern food business has become increasingly global and complex, challenging risks of fraud have arisen in various food industries.

As an important nutritional source for many types of diets, milk and milk products are in increasing demand in various parts of the globe. Milk frauds including dilution, substitution and addition of milk ingredients have been widely reported all over the world, but especially in developing countries (Afzal, Mahmood, Hussain, & Akhtar, 2011; Kandpal, Srivastava, & Negi, 2012; Shaikh, Soomro, Sheikh, Khaskheli, & Marri, 2013). United States Pharmacopeia has set up a database to record worldwide food fraud incidence up to date. Moore, Spink, and Lipp (2012) analysed the records in the database from 1980 to 2010 and indicated that the milk group is the second most commonly listed product group of the scholarly records dataset and the fifth on the list regarding the media and other records datasets. European Union developed the Rapid Alert System for Food and Feed (RASFF) to provide information on detected risks in food chains in EU countries. Based on RASFF, 16 notifications regarding milk and milk products from the Netherlands appeared in the past five years, two of those are related to fraud issues (one is about mislabelling and the other concerns illegal import).

Many analytical techniques have been used to develop milk authentication methods, such as high performance liquid chromatography-mass spectroscopy (Lu, Liu, Gao, Lv, & Yu, 2017), gas chromatography-mass spectroscopy (Li, Qi, & Shi, 2009), visible/near infrared spectroscopy (Hsieh, Hung, & Kuo, 2011; Liu, Parra, et al., 2018), mid-infrared spectroscopy (de Carvalho et al., 2015; Santos, Pereira-Filho, & Rodriguez-Saona, 2013), Raman spectroscopy (Nieuwoudt, Holroyd, McGoverin, Simpson, & Williams, 2016; Ullah et al., 2017) and proton transfer reaction mass spectrometry (Liu, Koot, Hettinga, de Jong, & van Ruth, 2018). Detection helps to limit further damage. However, even when detected, the fraud has already occurred and usually the foods have been distributed, and other actors in the chain and consumers have been deceived. Prevention on the other hand will be more effective in the longer term. It requires, however, understanding of the extent of the vulnerability to fraud and the factors contributing to this vulnerability (van Ruth, Huisman, & Luning, 2017). A food fraud vulnerability assessment (FFVA) is useful to explore vulnerability since it can identify areas in the food chain where vulnerabilities might exist and pinpoint key drivers and enablers (Spink, Moyer, & Whelan, 2016).The vulnerability to fraud is defined by three key elements:

opportunities, motivations, and control measures according to the food fraud concept of van Ruth et al. (2017). These elements can be subdivided into technical opportunities, opportunities in time and place, economic drivers, culture and behaviour, technical control measures, and managerial control measures. Based on the concept above, a practical FFVA tool of 50 questions was developed (SSAFE, 2015), and made available as a free downloadable app (PwC, 2017).

Fraud vulnerabilities in some other food supply chains were assessed with the FFVA tool previously (Bindt, 2016; Silvis, van Ruth, van der Fels-Klerx, & Luning, 2017; van Ruth, Luning, Silvis, Yang, & Huisman, 2018). The aim of the current study is to get insights in the general fraud vulnerability and underlying fraud factors of the milk supply chain in the Netherlands, as well as in the similarities and diversities between tier groups (farmers, processors, retailers). Moreover, fraud vulnerability at primary production level in different production systems was studied, i.e. for organic farms, green intermediate 'pasture' farms and conventional farms. Fraud vulnerabilities were assessed by the SSAFE FFVA tool, which was slightly adapted for the dairy chain.

2.2. Material and methods

2.2.1. Interviewed businesses

The study aimed to develop an overall insight into the fraud vulnerability of the Dutch dairy supply chain, including the tier groups farmers, processors and retailers. The main types of dairy farms in the Netherlands are conventional farms and pasture farms. They comprise more than 90% of all Dutch dairy farms. The pasture farm is a form of green agriculture specific to some countries. In the Netherlands, cows in pasture farms should stay outdoors at least 6 hours per day and 120 days per year (Liu, Koot, et al., 2018). The remaining group of farms is the group of organic farms. The organic products are gaining more and more attention from both the public and fraudsters, because of the specific way of production and its relatively high value compared to the other types of milk. Therefore we consider organic farm as the third group of farmers that will be assessed. In this study, 30 dairy farms (10 conventional farms, 10 pasture farms and 10 organic farms), 4 dairy processors (1 conventional and pasture processor, 1 organic processor and 2 all mixed processors; i.e. processors covering > 85% of the milk production in the Netherlands) and the 4 prime retailers in the Netherlands were assessed. The selection of the (locations of) the farms was carried out in such a manner that geographic and socio-economic scattering was maximized.

2.2.2. Adaptation of the SSAFE FFVA tool to the milk supply chain

The SSAFE FFVA tool was applied in the study. To make it more suitable to all tiers in the dairy supply chain, some minor changes were made. The questionnaires for the dairy farmers, processors and retailers differed slightly. Question 6 and question 7 in the original FFVA tool (SSAFE, 2015) were deleted since counterfeiting was not considered relevant for the dairy supply chain. This resulted in 48 questions overall and the relative factors are listed

in Table 2.1. In the questionnaire for the dairy farmers, questions about the direct supplier (Q18-22, Q39-42) were changed into corresponding questions about the direct customer, and questions about the raw material (Q2-3, Q30-31) were omitted as feed was not considered in this study. In the questionnaire for the retailers, the questions about their raw materials (because raw and final product is same), production activities and their direct customers (Q2-3, Q6, Q25, Q30-31) were omitted, since they are not applicable.

Table 2.1. The three key elements and 48 fraud factors in fraud vulnerability assessment.

Key elements	Fraud factors and their numbering
Opportunities	1. Ease of adulteration
	2. Availability of technology and knowledge for adulteration of the incoming milk (processors only)
	3. Detectability of fraud in incoming milk (processors only)
	4. Availability of technology and knowledge for adulteration of the final milk product
	5. Detectability of fraud in the final milk product
	6. Accessibility to production activities (N.A. for retailers)
	7. Transparency of the chain network
	8. Relationships within the supply chain
	9. Historical evidence of milk fraud
Motivations	10. Supply and price of milk
	11. Valuable components and attributes
	12. Economic health of the own company
	13. Business strategy of the own company
	14. Ethical business culture of the own company
	15. Previous irregularities of the own company
	16. Corruption level of the country in which the own company is active
	17. Financial pressure imposed by the company on the supplier (for the farms, financial pressure from the supplier)
	18. Supplier's (for the farmers, Customer's) economic health
	19. Supplier's (for the farmers, Customer's) business strategy
	20. Supplier's (for the farmers, Customer's) ethical business culture
	21. Supplier's (for the farmers, Customer's) previous irregularities
	22. Victimization of the supplier (for the farmers, Customer)
	23. Corruption level of the country in which the supplier/customer is active
	24. Economic health of the sector
	25. Customer's previous irregularities (N.A. for farmers and retailers)
	26. Sector ethical business culture
	27. Historical evidence of milk fraud, within sector
	28. Level of competition in sector
	29. Price asymmetries

Key elements	Fraud factors and their numbering
Control measures	30. Specificity and accuracy of the fraud monitoring system in place for incoming milk in the own company (processors only)
	31. Systematics and autonomy of verification of the fraud monitoring system for incoming milk in the own company (processors only)
	32. Specificity and accuracy of the fraud monitoring system in place for the milk product in the own company
	33. Verification of fraud monitoring system for the milk product in the own company
	34. Accuracy of the information system wrt mass balance control in the own company
	35. Extensiveness of the tracing and tracking system in the own company
	36. Application of integrity screening of employees in the own company
	37. Strictness of the ethical code of conduct in the own company
	38. Support of a whistle blowing system in the own company
	39. Contractual requirements with supplier in the own company (for the farmers, customers)
	40. Specificity and accuracy of the supplier's (for the farmers, customer) fraud monitoring system
	41. Accuracy of the supplier's (for the farmers, customer) information system wrt mass balance control
	42. Extensiveness of the supplier's (for the farmers, customer) tracking and tracing system
	43. Social control and transparency across the chain network
	44. Established guidance for fraud prevention and control in the sector
	45. Specificity of the national food policy
	46. Strictness of law enforcement in the local chain
	47. Strictness of law enforcement in the international chain
	48. Availability of a fraud contingency plan

2.2.3. Data collection

All the interviews with the farmers and retailers were conducted face-to-face. The processor surveys were conducted by face-to-face and phone interviews. Prior to the data collection, interviewers were trained and instructed on how to carry out the survey. All questions were translated into Dutch and sent to the respondents in advance. In case of the face-to-face interviews, some were conducted in English and some in Dutch. The interviewer asked the 48 questions and explained the questions if necessary. Three descriptions (without vulnerability scales), which were present in the original SSAFE FFVA tool, were provided with all 48 questions, and the respondents chose one of the optional answers to each question based on their knowledge and experience. All interviews lasted for 1.5-2 hours and were recorded with the permission of the interviewees.

2.2.4. Data analysis

A three-scale score system was adopted in the assessment. The descriptions to each question reflected a low, medium and high vulnerability level of the related fraud factor. For the fraud factors regarding opportunities and motivations, the descriptions were transformed into score 1, 2 and 3, which represented low, medium and high vulnerability level respectively. For the fraud factors associated with control measures, score 1, 2 and 3 represented low, medium and high adequacy of control measures respectively, which corresponded to high, medium and low vulnerability level. Some fraud factors were not available for the farmers or the retailers as indicated in section 2.2.2, they were left blank in this occasion.

For the compilation of the overall results of the assessment, the results of the three tier groups were balanced to allow them to contribute to the same extent: the weighted frequency of provided answers of each question (F_i) was determined by the following formula, Eq. (2.1)

$$F_i = \frac{1}{3}\sum_j {x_{ij}}\big/{n_j} \qquad\qquad \text{Eq. (2.1)}$$

Where F_i is the frequency of score i ($i = 1, 2, 3$), x_{ij} is the number of observations which get score i in group j (j = farmers, processors, retailers), n_j is the total number of observations in group j. The scores with highest F_i for the common fraud factors were used to present the overall results of the assessment.

The results of FFVA were subjected to multiple correspondence analysis (MCA) to investigate the association between groups, which was performed using R 3.4 (R Foundation for Statistical Computing, Vienna, Austria).

2.3. Results and discussion

The overall fraud vulnerability profiles and those of the tier groups in the Dutch milk supply chain are discussed based on the three key elements, opportunities, motivations and control measures in the following paragraphs. The key fraud factors are identified, and the similarities and differences in fraud vulnerability factors between tier groups assessed. Moreover, the vulnerability profiles of the three types of farms varying in production management are compared. The raw data of all interviews are attached as supplementary material.

2.3.1. Overall food fraud vulnerability data for the Dutch milk supply chain

For an overview of the fraud vulnerability and associated fraud factors, the weighted frequencies across all interviews for all common fraud factors were calculated (Fig. 2.1), and the modes for the three fraud elements for the overall supply chain are summarized in Fig. 2.2.
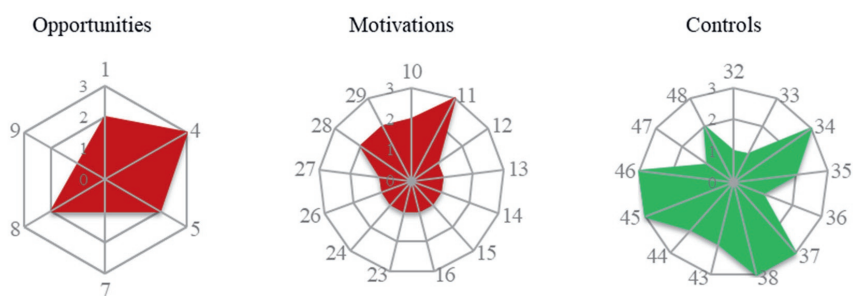
Fig. 2.1. Radar charts of scores with highest weighted frequency for common fraud factors in the fraud vulnerability assessment over all respondents, i.e. for opportunities (fraud factor 1-9), motivations (fraud factor 10-29) and control measures (fraud factor 32-48). Fraud factor numbers are listed in Table 2.1. Questions not applicable to all the tier groups are not shown.
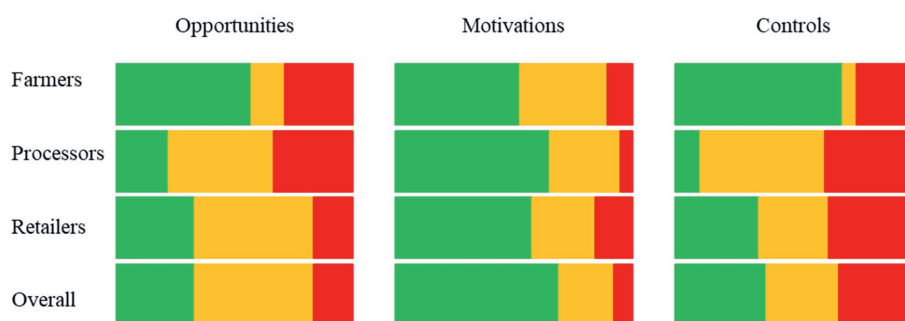
Fig. 2.2. The percentages of low, medium and high vulnerability modes of available individual factors from the vulnerability assessment, for the three fraud elements (opportunities, motivations and controls) and tier groups. The percentages of low, medium and high vulnerability of weighted modes are used for the overall supply chain. The low, medium and high vulnerability portions are coloured green, orange and red, respectively.

Opportunities. The radar chart on opportunity-related fraud factors (Fig. 2.1) shows that the fraud factors 1, 4, 5 and 8 are scored medium-high. The ease to adulterate milk (fraud factor 1) together with the availability of required technology for adulteration (fraud factor 4), and the general problem of detectability (fraud factor 5), contributes substantially to the general fraud vulnerability. The two factors, transparency of the supply chain (fraud factors 7) and historical evidence (fraud factor 9) are considered to contribute less to the fraud vulnerability. Overall, the opportunities are rated medium risk for the milk chain (Fig. 2.2).

Motivations. In the key element's group of motivations (Fig. 2.1), the valuable components and attributes (fraud factor 11) is rated high risk. According to the participants, the milk price is largely determined by its valuable components such as fat and protein content, and the special attributes such as organic or A2 type milk. The fluctuation of the supply and milk price, the competition level and price asymmetries (fraud factors 10, 28-29) were rated medium risk. The competition across the dairy industry might make it difficult for the chain actors to reach their financial goals, and the existing various price policies of milk and milk products

across countries might cause price fluctuation, both factors would motivate potential offenders to commit fraud (Huisman, 2013; van Ruth et al., 2017). All other nine factors were rated low risk. Overall the motivations-related fraud factors were assessed low to medium risk (Fig. 2.2).

Controls. The technical controls regarding fraud monitoring and its verification system, and contingency plan (fraud factor 32-33, 35, 48) are rated low-medium adequacy (Fig. 2.1), i.e. their absence contributes to the perceived fraud vulnerability in the supply chain. The lack of internal managerial controls such as the integrity screening of employees (fraud factor 36), as well as the shortage of external control measures such as sector guidance or (inter)national law enforcement (fraud factors 43-44, 47) increased fraud vulnerability. Overall, the controls were assessed as medium risk (Fig. 2.2).

Summarizing the overall results: due to the vulnerability scores for the key elements opportunities, motivations and control measures, the Dutch milk supply chain is considered low to medium vulnerable to fraud. There are, however, large differences between and within groups, which will be discussed in the following paragraphs.

2.3.2. Food fraud vulnerability data of the three tier groups: farmers, processors and retailers

To obtain an overall picture of the three main tier groups in the milk chain, all respondents' answers were subjected to MCA, the first two dimensions of which are shown in Fig. 2.3. The scores show distinct clustering of the three tier group businesses. There is a relatively large gap between the farmers and the retailers, and the processors are in the middle. A range of fraud factors is associated with the gap between the farmers and retailers and will be discussed in detail in section 2.3.2.2.
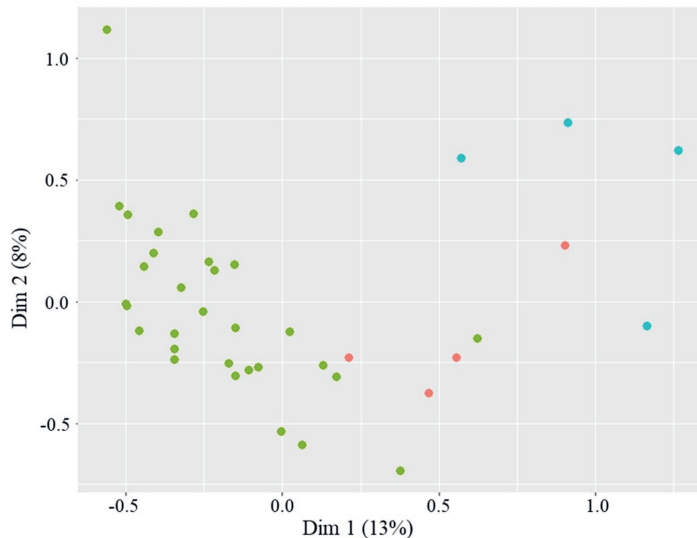
Fig. 2.3. Scores plot of the first two dimensions of multiple correspondent analysis on the common fraud factors of food fraud vulnerability assessment of farmers (green), processors (red) and retailers (blue). Fraud factors not applicable to all tier groups were omitted before analysis.

Table 2.2. Modes for individual fraud factors and tier groups from the fraud vulnerability assessments, ordered by fraud factor category. The low, medium and high vulnerability modes are coloured green, orange and red, and coded with H, M and L, respectively. The factors not applicable to particular tier groups are left blank and coded N.A. Fraud factor numbers are listed in Table 2.1.

Fraud factor category	No.	Farmers (n = 30)	Processors (n = 4)	Retailers (n = 4)
Technical opportunity	1	L	M	M
	2	N.A.	H	N.A.
	3	N.A.	H	N.A.
	4	M	H	H
	5	H	M	L
Opportunity in time and place	6	H	L	N.A.
	7	L	L	M
	8	L	M	M
	9	L	M	L
Economic drivers	10	H	M	L
	11	H	H	H
	12	L	L	L
	17	L	L	M
	18	L	M	L
	24	M	L	H
	28	M	M	H
	29	M	M	M
Culture and behaviour	13	L	L	L
	14	L	L	M
	15	L	L	L
	16	L	L	L
	19	L	L	L
	20	L	L	L
	21	L	L	L
	22	M	M	M
	23	M	L	L
	25	N.A.	M	N.A.
	26	M	L	M
	27	M	L	L
Technical controls	30	N.A.	H	N.A.
	31	N.A.	H	N.A.
	32	L	H	H
	33	L	H	H
	34	L	L	H
	35	L	M	M
	40	L	H	L
	41	L	M	L
	42	L	M	L
	48	H	M	M
Managerial controls	36	H	H	L
	37	H	L	L
	38	H	M	L
	39	L	M	M
	43	M	M	M
	44	L	M	H
	45	L	H	M
	46	L	M	H
	47	L	M	H

To illustrate the similarities and differences in fraud factors between the tier groups, the mode for each fraud factor for each tier group was calculated and is presented in Table 2.2, three key elements are sub-divided into six categories: technical opportunities, opportunities in time and place, economic drivers, culture and behaviour, technical controls and managerial controls. The high, medium and low vulnerability scores are coloured red, orange and green, respectively. In case of ties (equal frequencies), the colour of the higher vulnerability is shown in the table. The modes are summarized for the three key elements for each group in Fig. 2.2.

2.3.2.1. Similarities in relevant fraud factors among tier groups

The fraud factors presenting similar results (with the same mode score) across the tier groups are further detailed below (Table 2.2), i.e. a number of motivation-related and one control-related fraud factor. Fraud opportunity factors varied all in scores across the tiers.

Low vulnerability factors in common. All participants assigned low vulnerability scores to fraud factors associated with motivations (factor 12-13, 15-16, 19-21). They considered that they had long term financial targets and specific means to achieve their goals (fraud factor 12-13). None of the respondents was involved in the previous irregularities (fraud factor 15). However, it is noticeable that they estimated vulnerability resulting from their suppliers, supply chain and industry stage (fraud factor 22-27) higher than from themselves. One possibility is that there is a gap between the individuals and the entire sector with regard to the ethical business culture. Another possibility is that, for sensitive topics such as ethical behaviour, interviewees tend to stay out of the affairs, or to be overconfident about themselves, which might lead to bias of the results on such issues. This is in line with the alien conspiracy model and the techniques of neutralization. The former theory assumes that the crime is not a part of the own direct environment or shaped by the environment itself, but rather a problem of "outsiders" that threaten the environment (Kleemans, 2014). The rationalization technique "everybody else is doing it" has strong relation to the attitudes of business subculture. Most businessmen who believe their peers and competitors prefer to commit irregularities, are actually also doing that (Klenowski, 2012). As a result, the fraud vulnerability related to these factors might be concealed or underestimated. To decrease such bias, it is more objective to combine the result of self-assessment and assessment from the customer in terms of those factors.

Medium vulnerability factors in common. Victimization, price asymmetries and social control (fraud factor 22, 29, 43), are rated medium risk by all tier groups. Most of the respondents had no or quite limited information regarding whether their suppliers or customers had been the victim of food fraud, which leads the factor victimization (fraud factor 22) to receive the medium risk score. The difference of price policy of milk and milk product between some countries results in the medium vulnerability score for price asymmetries (fraud factor 29). The participants considered the milk supply chain has some degree of self-regulation, but the communication depends on the individual companies, and only serious issues would be discussed (fraud factor 43). Limited information exchange on unethical conducts is present in

the sector and therefore social control is rated medium risk by all groups.

High vulnerability factor in common. It has to be noticed that all three tier groups believed valuable components/attributes (fraud factor 11) as a high risk fraud factor. Generally, the milk price at farm level depends on its composition and quality. In the Netherlands, farmers are paid a lower price if their milk fails to meet the requirements, and they can get a bonus if they provide high quality milk continuously (Flores-Miyamoto, Reij, & Velthuis, 2014). The situation is similar in the U.S., where four classes of the farm-gate milk price are applied to set minimum prices for what the processors should pay dairy farmers (Bailey, 2005). Not only the valuable components, but some special attributes such as geographical indication or organic production also adds value to the milk (Caira, Pinto, Nicolai, Chianese, & Addeo, 2016; Luykx & van Ruth, 2008; Manca et al., 2001; Pillonel et al., 2002; Sharma & Paradakar, 2010). In some countries, a premium price is paid for special traits such as local production and grazing system (Tempesta & Vecchiato, 2013).

2.3.2.2. Differences in relevant fraud factors between tier groups

The fraud factors presenting diverse results (i.e. showing different modes) across the three tier groups are further discussed below. The three tier groups varied largely in assessing fraud opportunities and control measures. A few motivation factors were rated differently.

Opportunities. Despite some differences, technical opportunities (fraud factors 1-5) are rated medium to high risk. Only simple knowledge or basic technologies are required to adulterate milk (fraud factor 2, 4). Raw milk and liquid milk products are materials with a complex composition, which in general are easier to manipulate. This is consistent with the criteria that liquid material, which can be freely accessed to all adulterants, is the easier physical state to be adulterated (Jack, 2015). The farmers and processors considered advanced analysis or test was required to confirm milk authenticity (fraud factor 5). Moyer, DeVries, and Spink (2017) indicate that it is very challenging or even impossible to confirm the authenticity of milk product, considering the milk products contain thousands of molecules that all react or interact under certain conditions. The retailers considered the detection of milk adulteration was easy. This might be due to the retailers judged of the commercial milk only, which was standardized and homogenous or due to lack of technical knowledge. The processing activities (fraud factor 6) of the dairy processors are continuous and access extremely strictly limited. On the contrary, most of the Dutch dairy farms are quite open, both to their employees and the public. Only few of the farmer respondents mentioned that they locked the milk tank storage off-hours. Although the lack of physical safeguards is not directly linked to food fraud vulnerability, previous research points out that it will increase the opportunity for external threat (van Ruth et al., 2017). The farmers considered the supply chain relationship (fraud factor 8) as integrated and well-coordinated, but the processors and retailers had a different opinion. According to them, the supply chain is not fully transparent.

Motivations. A large difference in rating among the tier groups exists for economic driver related factors. The farmers, processors and retailers rated supply and price of milk (fraud factor 10) as high, medium and low risk, respectively. The milk supply and price was considered fluctuating by the farmers in the current study. A number of factors determine the milk price at different points of the supply chain (GAO, 2004). Fluctuation of the milk price in the Netherlands was observed in the past decades, but the volatility has increased recently with the change of the EU subsidy programmes. The causes behind farm milk price volatility generally depend on weather patterns, which are becoming more extreme and less predictable. They have a larger influence on those whose cows are primarily fed on grass (Jongeneel & Berkum, 2015; Zuivel NL, 2016). The international trade activities such as the increased import level of dairy products from the world market to the Dutch or EU market, and the Russian trade ban on milk products, were the economic reasons of milk price fluctuation in the past few years. In April 2015, the European Union milk quota system was abolished in response to the increasing global demand for milk (EC, 2015). These uncertainties made the milk price more volatile. The farmers and processors rated their economic health as being in good condition (fraud factor 12), but the retailers considered the liquid milk market declining due to the decrease in sales (fraud factor 24). One has to consider though that the retailers considered liquid consumption milk only. However, the farmers produce liquid milk not only for direct consumption, but most will end up as milk powders, butter and cheese (Zuivel NL, 2016). Unlike the farmers, the processors were less confident about their suppliers' economic health (fraud factor 18). This is also reflected by the figures of Dutch dairy sector. While the production value of milk processing remained stable, the income development of dairy farms decreased largely from 2014 (Zuivel NL, 2016). This financial situation affects fraud vulnerability. Another key driver which was rated differently is competition level (fraud factor 28). Despite the processors and farmers considered the competition within dairy sector was moderate, the retailers rated competition as high risk. Intense competition could drive food fraud, especially if there are organizations that have better economies of scale in the supply chain (Manning, 2016).

Controls. All control measures are rated differently except for the social controls (fraud factor 43). The technical control measures are rated as highly adequate by the farmers, however, are considered the opposite by the processors and retailers. The farmers were short of internal managerial controls (fraud factor 36-38), whereas the processors and retailer considered managerial controls in the wider environment as inadequate (fraud factor 44-47), i.e. from local dairy sector, national authorities or global organizations.

The processors and retailers assigned low adequacy scores to the fraud monitoring system and its verification (fraud factors 32-33), indicating their sampling plans were mainly for food safety and quality, but no fraud monitoring tasks or verifications were carried out. Elliott (2014) pointed out that current food safety audits are hardly designed to assure the authenticity of food products or to identify fraudulent practices, which is in line with the viewpoint of the processors and retailers in the Dutch dairy chain. The farmers, however, assigned high adequacy to fraud monitoring system and its verification (fraud factor 32-33). "Milk tank drivers are responsible for the rapid on-site test, and the dedicated laboratory then carries out the analyses." was

commented by most of the respondent farmers on question 32. It should be noticed that the current monitoring system are performed for payment decision, rather than fraud prevention (Flores-Miyamoto et al., 2014). The verification of the monitoring system on dairy farms is in the form of audits. Every farm in the Netherlands is audited every two years on numerous standards (van Asseldonk & Velthuis, 2014). The audits commonly focus on safety and quality issues, but the farmers consider that it might work for fraud issues as well. The processors and retailers have some contingency plans only for food safety and quality issues, whereas few farmers had anything in place (fraud factor 48). Contingency plans are designed to deal with emergencies and can be deemed as a last resort in protection of business from fraud threats (van Ruth et al., 2017).

A gap existed when the farmer and processor assessed each other in regard to the technical controls. While the farmers perceived the technical measures, mass balance control and trace and tracking system (fraud factor 34-35), in their own farms as highly adequate, the processors rated those of their suppliers (fraud factor 41-42) as low-medium adequate. The retailers rated these technical measures applied by their suppliers (fraud factor 41-42) as highly adequate, but the processors rated those within their businesses (fraud factor 34-35) as moderately adequate for fraud monitoring. This gap might due to the limited information exchanged between tier groups.

The managerial control measures showed considerable differences between the three tiers (Table 2.2). While the farmers were lacking various internal managerial control measures (fraud factor 36-38), the retailers had relatively sufficient ones established such as standardized integrity screening methods for employees, well-established whistle blowing system and detailed written code of ethical conduct. The managerial control measures in the processors were in the middle between the farmers and retailers, at medium adequacy level. The reason of the lack of measures such as "integrity screening of employees", "ethical code of conduct" and "whistle blowing" (fraud factors 36-38) on farm level might be that many farms are family businesses, and they have no or only a few employees (the average number of full time employees in the farms is 2.9 in this study). The farmers doubted, during the interviews, if the integrity screening and whistle blowing were necessary for them. Despite the fact that the processors and retailers, as large scale companies, had managerial controls to some extent in place, they were also perceiving a general lack of sufficient control measures. This lack of controls was primarily related to the wider supply chain, such as lack of guidance on fraud prevention, or available regulations and adequate enforcement (fraud factor 44-46). The farmers, however, rated guidance and regulation/law (enforcement) to prevent food fraud (fraud factors 44-47) sufficient. The retailers considered that the government should be responsible for setting the minimum standards in food safety, and they thought this was even more urgent for food fraud issues. Most retailers not only require their suppliers to meet general food safety standards and be certified, they have also drawn up standards on their own to decrease or eliminate food fraud risks resulting from their suppliers, and to ensure the integrity of the food product.

2.3.2.3. Overall vulnerability of tiers

The number of low-medium-high risk modes for the three key elements are summarized for each tier group in Fig. 2.2. For the opportunities, the farmers score lower vulnerability than the processors and retailers. For the motivations, similar results for the three groups are obtained. The controls are rated differently for the three groups, with farmers providing lower vulnerability scores across the board. Thus, the processors and retailers consider themselves more vulnerable to milk fraud than the farmers group.

2.3.3. Differences between farms of different management systems

To obtain an overview of the similarities and dissimilarities of the three farm groups (conventional, pasture and organic farmers), the answers from all farmers were subjected to MCA, the first two dimensions of which are presented in Fig. 2.4. The MCA plot shows that the organic and non-organic farmers are clearly separated, whereas the conventional and pasture farmers present a more mixed pattern. To illustrate the diversity on fraud factors between the organic and non-organic farmers, the mode for each fraud factor for each type of farmers group was calculated and presented in Table 2.3. The high, medium and low vulnerability scores are coloured red, orange and green, respectively. In case of ties (equal frequencies), the colour of the higher vulnerability is shown in the table. The percentages of the low, medium and high vulnerability mode of available individual factors, for the farmer groups and fraud elements (opportunities, motivations and controls) are summarized in Fig. 2.5.
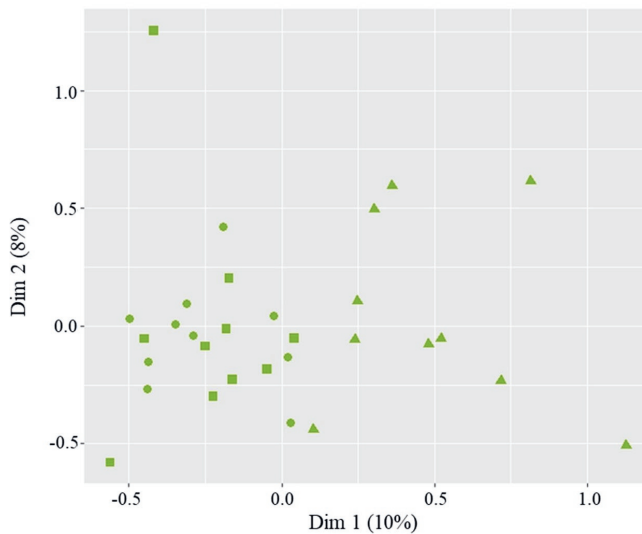

Fig. 2.4. Scores plot of the first two dimensions of multiple correspondent analysis on the fraud vulnerability assessment results of the organic farmers (triangle), pasture farmers (square) and conventional farmers (round).

Table 2.3. Modes for individual fraud factors and three types of farmers from the fraud vulnerability assessments, ordered by fraud factor category. The low, medium and high vulnerability modes are coloured green, orange and red, and coded with H, M and L, respectively. The factors not applicable to the farmer group are omitted. Fraud factor numbers are listed in Table 2.1.

Fraud factor category	No.	Conventional farmers (*n* = 10)	Pasture farmers (*n* = 10)	Organic farmers (*n* = 10)
Technical opportunity	1	L	M	H
	4	M	H	M
	5	M	H	H
Opportunity in time and place	6	H	M	H
	7	L	M	L
	8	M	L	L
	9	L	L	M
Economic drivers	10	H	M	L
	11	H	H	H
	12	M	L	L
	17	L	L	L
	18	L	L	L
	24	M	M	L
	28	M	M	L
	29	M	H	M
Culture and behaviour	13	L	L	L
	14	L	L	L
	15	L	L	L
	16	L	L	L
	19	L	L	L
	20	L	L	L
	21	L	L	L
	22	M	M	L
	23	M	M	L
	26	M	M	M
	27	M	M	M
Technical control measure	32	L	L	M
	33	L	L	M
	34	L	L	L
	35	L	L	L
	40	L	L	L
	41	L	L	M
	42	L	L	M
	48	H	H	M
Managerial control measures	36	H	H	H
	37	H	H	H
	38	H	H	H
	39	L	L	L
	43	L	M	M
	44	L	M	M
	45	L	L	L
	46	L	L	L
	47	M	L	L

Fig. 2.5. The percentages of low, medium and high vulnerability mode of available individual factors from the vulnerability assessment, for the three fraud elements (opportunities, motivations and controls) and three types of farmers. The low, medium and high vulnerability portions are coloured green, orange and red, respectively.

The fraud factors with different ratings between the organic and non-organic farmers are further discussed below. Different vulnerability scores were assigned to the opportunity factors ease of adulteration and historical evidence (fraud factor 1, 9). Motivations-related factors supply and price, victimization of customers, corruption level of the countries in which customer is active, the sector's economic health and price asymmetries (fraud factor 10, 22-24, 28) show different results for the three groups. The controls with regard to fraud monitoring system and verification, customer's mass balance control and trace and tracking system, and contingency plan (fraud factor 32-33, 41-42, 48) show also different vulnerability scores.

Opportunities. Although all farmer groups believed the technical opportunities (fraud factor 1, 4-5) contributed considerably to fraud vulnerability, the organic farmers assumed the organic milk was even easier to be adulterated. "It is quite easy to mix any kind of milk with organic milk to commit fraud, which is also very difficult to detect" commented by quite a few organic farmers. The non-organic farmers believed no milk frauds happened in the past (fraud factor 9), which is surprising, whereas the organic farmers had no solid information. Overall, the fraud opportunities are perceived as slightly higher by the organic farmers compared with the rest, which makes sense (Fig. 2.3).

Motivations. Supply and price of milk (fraud factor 10) is graded differently by the three farmer groups. While the organic milk price fluctuated slightly and the supply of organic milk was sufficient, the conventional milk supply and price was considered more volatile. "The demand of organic milk increased steadily, but the price remained relatively stable." was commented by most organic farmers. This is consistent with the data on milk price development in the Dutch agricultural sector (Jongeneel & Berkum, 2015). The conventional farmers considered sector economic health (fraud factor 24) was not quite profitable and the market is stable, which contributed moderately to fraud vulnerability. On the contrary, the organic farmers believed they are in a growing market. According to the report from IFOAM organics

international, the organic milk yield in European Union has almost doubled since 2007 to meet the increasing demand, constituting around 3 percentage of EU milk production in 2016 (Willer & Lernoud, 2018). The competition level (fraud factor 28) in the non-organic farms was rated medium risk, but in the organic farms as low risk. The organic farmers thought their suppliers and customers has not been involved in fraud incidents in the past, and they are active in a country with low corruption levels (fraud factor 22-23). The non-organic farmers had no confirmed information about their customers' victimization. It should be noticed in the comparison between the organic and non-organic farms that the impact of the price premium of the organic milk was slightly weaken, because the related fraud factor, valuable component and special attribute, reflected the added value of both organic and high protein and fat content. Overall fraud motivations in the organic farms appear lower than the motivations in the non-organic farms (Fig. 2.5).

Controls. The organic and non-organic farmer groups rated a series of technical controls differently. The organic farmers considered that a fraud monitoring system and its verification system (fraud factor 32-33) were moderately adequate. Although the non-organic farmers assigned high adequacy to those factors, they pointed out that the measures aimed mainly at food safety and quality issues, which was discussed in 3.2.2. The organic farmers had very limited information about their customers' mass balance control and track and trace system (fraud factor 41-42), and rated those factors medium adequate. The non-organic farmers held positive opinions towards their customers with respect to their control measures. Most of the conventional and pasture farmers have no contingency plans (fraud factors 48). The organic farmers indicated that they had some contingency measures to deal with food safety and quality issues, but not specifically for food fraud. Overall, lack or lower adequacy of controls add more vulnerability at farm level to the organic milk production than to the conventional production.

The organic farms show higher vulnerability due to opportunities, lower vulnerability in regard to motivations, and higher vulnerability due to controls than non-organic farms. Altogether they appear slightly more vulnerable than their non-organic counterparts.

2.3.4. Methodology consideration

The SSAFE FFVA tool was initially designed for self-assessment for the food businesses to identify the fraud vulnerability in food ingredient, product or food company. The respondents' qualification including educational level, working experience, knowledge about food fraud, etc. may have impact on their answers in this study. Although all the respondents have been well informed about the aim of the survey, this formal approach (interview from the university) may increase the interviewees' concern or heighten their evaluation of the risk. These factors need to be noted and assessed in the future research.

2.4. Conclusions and outlook

From the assessments is concluded that, overall, the Dutch milk supply chain is perceived as low to medium vulnerable to food fraud according to the interviewed farmers, processors

and retailers. The processors and retailers are more vulnerable to milk fraud than the farmers. The three tier groups showed similarities with regard to fraud motivations, but large differences regarding vulnerability resulting from fraud opportunities and controls. Furthermore, the organic farmers appeared to be slightly more vulnerable than their non-organic counterparts.

Based on the current study, it is suggested that a milk price monitoring system to discover any anomalous price fluctuation aligned with advanced fraud monitoring systems and adequate detection methods will help individual businesses to reduce vulnerability. Furthermore, generally food fraud mitigation and prevention can be enhanced by more industry guidelines, and stricter fraud dedicated national food policies and fraud enforcement. The current study focused on vulnerability, i.e. from the viewpoint of companies becoming a victim. Future change of perspective towards offenders would provide additional knowledge, especially for authorities and certification bodies.
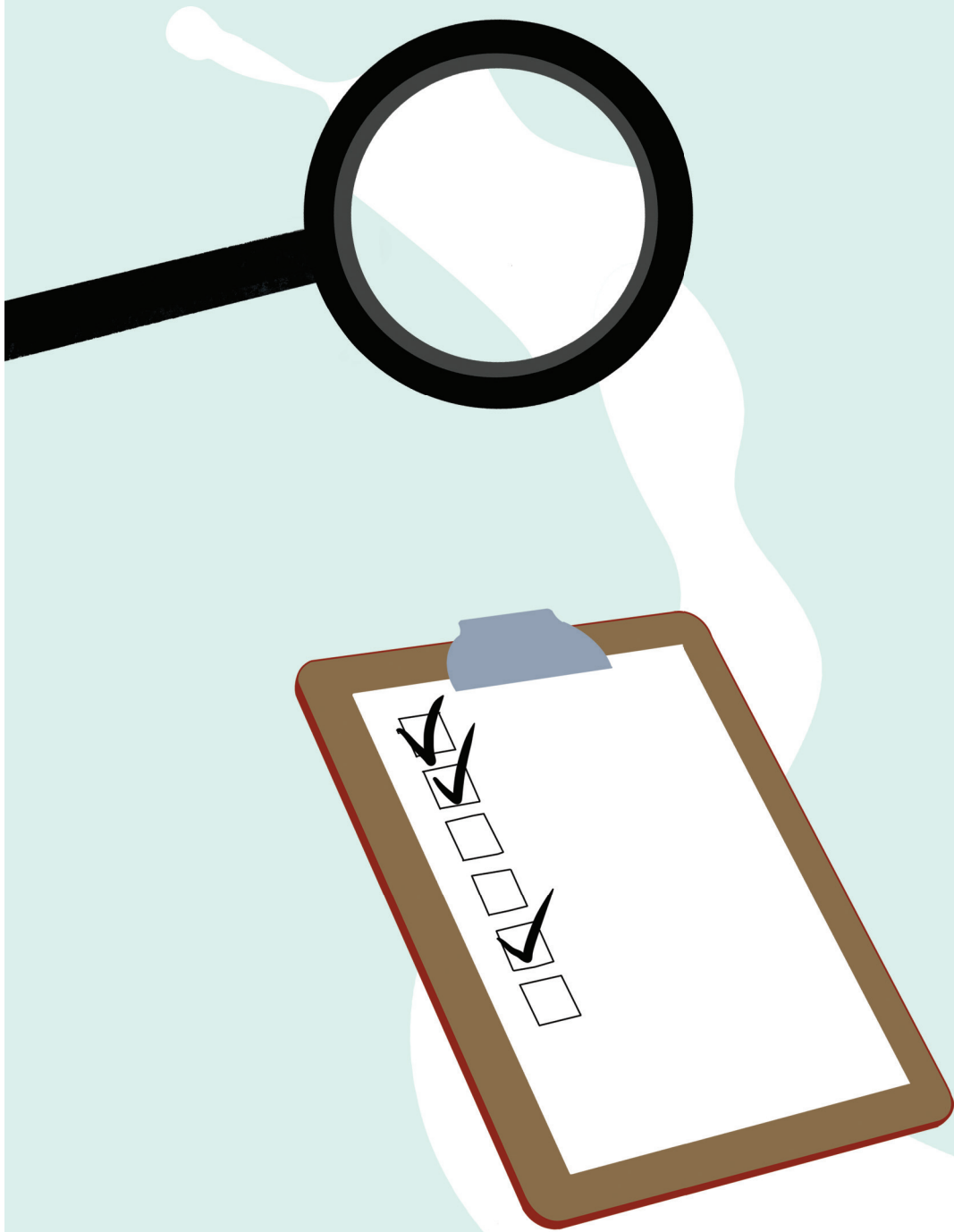
Acknowledgements

2

References

Afzal, A., Mahmood, M., Hussain, I., & Akhtar, M. (2011). Adulteration and microbiological quality of milk (a review). *Pakistan Journal of Nutrition, 10*(12), 1195-1202.

Bailey, K. (2005). Dairy risk-management education: Factors that affect U.S. farm-gate milk prices. Retrieved from http://extension.psu.edu/publications/ua412/view.

Bindt, V. (2016). *Costs and Benefits of the Food Fraud Vulnerability Assessment in the Dutch Food Supply Chain*. MSc thesis. Wageningen University and Research, the Netherlands. Retrieved from https://edepot.wur.nl/390258.

Caira, S., Pinto, G., Nicolai, M. A., Chianese, L., & Addeo, F. (2016). Simultaneously tracing the geographical origin and presence of bovine milk in Italian water buffalo Mozzarella cheese using MALDI-TOF data of casein signature peptides. *Analytical and Bioanalytical Chemistry, 408*(20), 5609-5621.

de Carvalho, B. M. A., de Carvalho, L. M., dos Reis Coimbra, J. S., Minim, L. A., de Souza Barcellos, E., da Silva Júnior, W. F., Detmann, E., & de Carvalho, G. G. P. (2015). Rapid detection of whey in milk powder samples by spectrophotometric and multivariate calibration. *Food Chemistry, 174*, 1-7.

European Commision. (2015). The end of milk quotasotas. Retrieved from https://ec.europa.eu/agriculture/milk-quota-end_en.

Elliott, C. (2014). *Elliott Review into the Integrity and Assurance of Food Supply Networks*. London: Department for Environment, Food & Rural Affairs.

Flores-Miyamoto, A., Reij, M. W., & Velthuis, A. G. J. (2014). Do farm audits improve milk quality? *Journal of Dairy Science, 97*(1), 1-9.

GAO. (2004). Dairy industry information on milk prices, factors affecting prices, and dairy policy options. Retrieved from http://www.gao.gov/products/GAO-05-50.

Hsieh, C. L., Hung, C. Y., & Kuo, C. Y. (2011). Quantization of adulteration ratio of raw cow milk by least squares support vector machines (LS-SVM) and visible/near infrared spectroscopy. In *Engineering Applications of Neural Networks*. Berlin, Germany: Springer.

Huisman, W. (2013). Criminogenic organizational properties and dynamics. In *The Oxford Handbook of White-Collar Crime*. Oxford: Oxford University Press.

Jack, L. (2015). *Risk Modelling of Food Fraud Motivation: 'NSF Fraud Protection Model' Intelligent Risk Model Scoping Project FS 246004: Final Report*. London: Food Standard Agency.

Jongeneel, R., & van Berkum, S. (2015). *What Will Happen after the EU Milk Quota System Expires in 2015? : An Assessment of the Dutch Dairy Sector*. LEI Wageningen Univeristy and Research, Netherlands. Retrieved from http://edepot.wur.nl/338308.

Kandpal, S. D., Srivastava, A. K., & Negi, K. S. (2012). Estimation of quality of raw milk (open & branded) by milk adulteration testing kit. *Indian Journal of Community Health, 24*(3), 188-192.

Kleemans, E. R. (2014). Theoretical perspectives on organized crime. In *Oxford Handbook of Organized Crime*. Oxford: Oxford University Press.

Klenowski, P. M. (2012). "Learning the good with the bad" are occupational white-collar offenders taught how to neutralize their crimes? *Criminal Justice Review, 37*(4), 461-477.

Li, J., Qi, H., & Shi, Y. (2009). Determination of melamine residues in milk products by zirconia hollow fiber sorptive microextraction and gas chromatography-mass spectrometry. *Journal of Chromatography A, 1216*(29), 5467-5471.

Liu, N., Koot, A., Hettinga, K., de Jong, J., & van Ruth, S. M. (2018). Portraying and tracing the impact of different production systems on the volatile organic compound composition of milk by PTR-(Quad)MS and PTR-(ToF)MS. *Food Chemistry, 239*(Supplement C), 201-207.

Liu, N., Parra, H. A., Pustjens, A., Hettinga, K., Mongondry, P., & van Ruth, S. M. (2018). Evaluation of portable near-infrared spectroscopy for organic milk authentication. *Talanta, 184,* 128-135.

Lu, W., Liu, J., Gao, B., Lv, X., & Yu, L. (2017). Technical note: Nontargeted detection of adulterated plant proteins in raw milk by UPLC-quadrupole time-of-flight mass spectrometric proteomics combined with chemometrics. *Journal of Dairy Science, 100*(9), 6980-6986.

Luykx, D. M. A. M., & van Ruth, S. M. (2008). An overview of analytical methods for determining the geographical origin of food products. *Food Chemistry, 107*(2), 897-911.

Manca, G., Camin, F., Coloru, G. C., Del Caro, A., Depentori, D., Franco, M. A., & Versini, G. (2001). Characterization of the geographical origin of pecorino sardo cheese by casein stable isotope (13C/12C and 15N/14N) ratios and free amino acid ratios. *Journal of Agricultural and Food Chemistry, 49*(3), 1404-1409.

Manning, L. (2016). Food fraud: Policy and food chain. *Current Opinion in Food Science, 10*, 16-21.

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.

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.

2

Nieuwoudt, M. K., Holroyd, S. E., McGoverin, C. M., Simpson, M. C., & Williams, D. E. (2016). Raman spectroscopy as an effective screening method for detecting adulteration of milk with small nitrogen-rich molecules and sucrose. *Journal of Dairy Science, 99*(4), 2520-2536.

Pillonel, L., Badertscher, R., Bütikofer, U., Casey, M., Dalla Torre, M., Lavanchy, P., Meyer, J., Tabacchi, R., & Bosset, J. (2002). Analytical methods for the determination of the geographic origin of Emmentaler cheese. Main framework of the project; chemical, biochemical, microbiological, colour and sensory analyses. *European Food Research and Technology, 215*(3), 260-267.

PwC. (2017). Food fraud vulnerability assessment. Retrieved from https://www.pwc.com/gx/en/services/food-supply-integrity-services/publications/food-fraud.html.

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.

Shaikh, N., Soomro, A., Sheikh, S., Khaskheli, M., & Marri, A. (2013). Detection of adulterants and their effect on the quality characteristics of market milk. *Pakistan Journal of Agricultural Engineering and Veterinary Sciences, 29*(2), 175-183.

Sharma, K., & Paradakar, M. (2010). The melamine adulteration scandal. *Food Security, 2*(1), 97-107.

Silvis, I. C. J., van Ruth, S. M., van der Fels-Klerx, H. J., & Luning, P. A. (2017). Assessment of food fraud vulnerability in the spices chain: An explorative study. *Food Control, 81*, 80-87.

Spink, J., & Moyer, D. C. (2011). Defining the public health threat of food fraud. *Journal of Food Science, 76*(9), R157-R163.

Spink, J., Moyer, D. C., & Whelan, P. (2016). The role of the public private partnership in food fraud prevention – includes implementing the strategy. *Current Opinion in Food Science, 10*, 68-75.

SSAFE. (2015). Food fraud vulnerability assessment tool. Retrieved from http://www.ssafe-food.org/our-projects/.

Tempesta, T., & Vecchiato, D. (2013). An analysis of the territorial factors affecting milk purchase in Italy. *Food Quality and Preference, 27*(1), 35-43.

Ullah, R., Khan, S., Ali, H., Bilal, M., Saleem, M., Mahmood, A., & Ahmed, M. C. J. R. S. R. (2017). Raman-spectroscopy-based differentiation between cow and buffalo milk. *Journal of Raman Spectroscopy, 48*(5), 692-696.

van Asseldonk, M. A. P. M., & Velthuis, A. G. J. (2014). Risk-based audit selection of dairy farms. *Journal of Dairy Science, 97*(2), 592-597.

van Ruth, S. M., Huisman, W., & Luning, P. A. (2017). Food fraud vulnerability and its key factors. *Trends in Food Science & Technology, 67*, 70-75.

van Ruth, S. M., Luning, P. A., Silvis, I. C. J., Yang, Y., & Huisman, W. (2018). Differences in fraud vulnerability in various food supply chains and their tiers. *Food Control, 84*, 375-381.

Willer, H., Lernoud, J., & Kemper, L. (2018). The world of organic agriculture 2018: Summary. In *The World of Organic Agriculture. Statistics and Emerging Trends 2018*, Bonn, Germany. Retrieved from https://riau.ac.ir/Files/b.pdf#page=28.

Zuivel NL (2016). Dutch dairy in figures 2016. Retrieved from http://www.zuivelnl.org/wp-content/uploads/2017/06/Dutch-dairy-in-figures-2016.pdf.

2

CHAPTER 3

The Chinese milk supply chain: A fraud perspective

Abstract

Food fraud has become a serious concern all over the world and especially in China. The melamine contaminated infant formula in 2008 has brought food fraud in the spotlight. This incident had grave consequences for the Chinese citizens as well as the Chinese milk industry. Fraud vulnerability assessments are the first step towards food fraud prevention and mitigation. To combat food fraud, one has to think like a criminal. In the current study, we determined the most vulnerable points in the Chinese milk supply chain, and examined the underlying causes. The fraud vulnerability perceived by 90 Chinese dairy farmers and 14 milk processors was evaluated with the SSAFE food fraud vulnerability assessment tool. Overall, actors perceived the milk supply chain as low to medium vulnerable to food fraud. Farmers appeared significantly more vulnerable than processors due to enhanced opportunities and motivations, and less adequate controls they perceived. Both geographical location of the farms and their size affected their perceived fraud vulnerability significantly.

Keywords

3.1. Introduction

Food fraud is generally considered the illicit deception for economic gain using food, food ingredients or food packing (Spink & Moyer, 2011), and has attracted increasing attention from the food industry and regulators. Food fraud threatens food integrity in the form of manipulation of food products, food processing, information recording systems and also personal integrity (Manning, 2016). The consequences of food fraud incidents are various, and it takes extensive time to recover from fraud incidents.

Dairy products are an important source of protein and calcium in the human diet, which also holds for China. Due to the growing population and wealth in China, its dairy production increased by more than 20% annually from 1997, and reached 35 million tons in 2007 (Qian, Guo, Guo, & Wu, 2011). The melamine incident in 2008, when melamine was added to apparently inflate the protein content of animal feed and milk, was a dreadful turn for the Chinese dairy industry. Eventually, the melamine ended up in infant formula. Melamine cyanuric acid, a salt, is formed in the presence of melamine and cyanuric acid. Since this salt does not dissolve easily and forms crystals, it can lead to bladder and kidney stones, and subsequent acute renal failure. Infants are a particularly vulnerable group, because their organs have yet to form fully and their nutrition is more restricted (Pei et al., 2011). The melamine contamination of infant formula caused illness of 300,000 individuals, hospitalisation of 50,000 infants and six deaths (Graham-Harrison, 2009). Besides the severe health price paid by the consumers of the products, the estimated financial loss of Chinese dairy industry in 2008 from the melamine incident was RMB 20 billion (Wang, 2009). The milk production in China levelled off after the melamine incident instead of showing further growth (Li, 2016). The government imposed severe penalties on the people involved in the incident, including execution of two persons and life imprisonment of four persons (Xiu & Klein, 2010). Nevertheless, countries around the world banned the import of Chinese dairy products for an extended period. The incident had a large impact on the trust of Chinese consumers themselves too, many of which still seek for import infant formula instead of locally produced products today, a decade after the incident.

The melamine incident was a very visible and representative example of the numerous food fraud incidents that emerged across the Chinese food production chains. With rapid economic growth, the living standard has risen across Chinese citizens and this had consequential effects on their food and diets (Lam, Remais, Fung, Xu, & Sun, 2013). Chinese consumers are becoming more aware of food quality and safety issues (Ouyang, 2011). Although there has been an increasing focus on food safety governance in the past decade (Zhu, Huang, & Manning, 2019), food fraud incidents kept emerging (Zhang & Xue, 2016).

In the more distant past, science and industry focused primarily on fraud detection and methodology for detection. In the past five years, there is also more attention paid to prevention (Cadieux, Goodridge, & Spink, 2019). The latter requires better understanding of what drives food fraud. Some conceptualised food fraud (Spink & Moyer, 2011; van Ruth, Huisman, &

Luning, 2017), others profiled food fraud incidents (Moore, Spink, & Lipp, 2012; Zhang & Xue, 2016). An important step towards fraud prevention is to understand the extent of the fraud vulnerability and determine which factors contribute to this vulnerability. Research has established that food fraud vulnerability could be initiated from both internal and external environment of a business, thus one needs to consider both the internal dark side as well as the external threats (van Ruth et al., 2017). A food fraud vulnerability assessment (FFVA) tool can help to identify the weaker spots in the chain network, and thereby provides information on the potential factors contributing to food fraud (van Ruth, Luning, Silvis, Yang, & Huisman, 2018). Previously, we evaluated the fraud vulnerability in the milk supply chain in the Netherlands (Yang et al., 2019).

Considering the previous milk fraud incidents in China and steps made towards improvements, it is of interest to examine the fraud vulnerability of actors in the Chinese milk supply chain today. In the current study we assessed the perceived fraud vulnerability by the main actors of milk supply chain, and underlying factors of Chinese milk businesses, 90 farmers and 14 processors, from the perspective of discerning the weakest spots. Thus, using the 'thinking like a criminal' approach (Levitt & Dubner, 2014), in term of food fraud vulnerabilities identification (Spink, 2019). We evaluated and compared the farmers and processors, but also examined the effect of geography and size of the farms. The perceived vulnerabilities of the Chinese actors were compared to those of their Dutch counterparts too.

3.2. Methodology

3.2.1. Participant recruitment

A two-step mixed method was adopted to collect data. First, a digital survey was conducted. Dairy farmers and processors were invited to participate in this survey via email and calls through the National Dairy Industry and Technology System. The first author presented at the annual meeting of National Dairy Industry and Technology System in 2017 and gave a brief introduction about the survey. The study focuses on fraud vulnerability of the milk supply chain, so only the farmers and processors involved in milk production were invited for the participation of this survey. A total of 90 dairy farmers and 14 milk processors returned completed questionnaire. Subsequently, interviews with Chinese milk experts ($n = 4$) were conducted (using the same questionnaire), to get further insights on these fraud factors. These experts consisted of two farmers, one processor from the industry and one professor from academia.

The survey respondents represent nationwide participants in the main milk production area in China. As shown in Table 3.1 and 3.2, the participants are scattered in the northern and eastern China, where more than 80% of raw milk and more than 75% processed milk are produced. Moreover, the dairy farmers from the north consisted around 80% of the farmer respondents in this study, which matches the fact that the northern China is the main raw milk production area (more than 70% raw milk is produced), (Table 3.1).

Table 3.1. Demographic characteristics of the farmers in the study, and the related regional data in China.

Variable	Item	Number of farms (percentage)	Regional data in China *		
			Distribution of raw milk production	Distribution of number of cows	Distribution of number of farms
Location (*n* = 90)	Central-north	54 (60%)	40%	33%	13%
	Northeast	10 (11%)	20%	17%	21%
	Northwest	10 (11%)	14%	13%	50%
	East	16 (18%)	12%	23%	3%
	Others	0 (0%)	14%	14%	13%
Size (*n* = 90)	Small (< 500 cows)	26 (28%)			
	Medium (500-1,000 cows)	32 (36%)			
	Large (> 1,000 cows)	32 (36%)			
Age of the respondents (*n* = 79) **	20-30	3 (4%)			
	30-40	19 (24%)			
	40-50	26 (33%)			
	50-60	24 (30%)			
	> 60	7 (9%)			
Number of farm employees (*n* = 82) **	1-10	12 (15%)			
	11-50	54 (66%)			
	51-100	10 (12%)			
	> 100	6 (7%)			

* Data retrieved from Ministry of Agricultural P. R. China (2016). ** Total participant number varies since some respondents did not provide the completed demographic information.

3.2.2. Adaptation of the SSAFE FFVA tool to the milk supply chain

A practical FFVA questionnaire, containing various factors contributing to food fraud, was developed previously (SSAFE, 2017). The original questionnaire contains 50 questions to evaluate the three key elements for the perpetration of food fraud, i.e. fraud opportunities, fraud motivations and controls. In other words: the possibility and willingness to offend by the availability of a suitable target to adulterate and the absence of retrains. The assessment of both internal and external environment of the business has been involved in the questionnaire. Only the economically motivated adulteration was considered with regard to milk fraud. The SSAFE questionnaire was adapted to dairy farms and processors for this study. Questions 6 and 7 in the original FFVA tool (SSAFE, 2017) were deleted because counterfeiting was not considered relevant for the milk production chain. As a result, 48 questions were used in the questionnaire (Table 3.3). In the questionnaire for the farmers, the questions (Q18-22, Q39-42) regarding the suppliers were changed into the corresponding ones about the customers, i.e. the processors. The questions about raw material (Q2-3, Q30-31) were omitted as feed was not considered in this study, and question 25 was removed because it replicated question 21 in this specific situation. This survey was conducted in form of self-assessment, and all the participants were required to answer the questionnaire according to their own milk business.

Table 3.2. Demographic characteristics of the processors in the study, and the related regional data in China.

Variable	Item	Number of processors (percentage)	Regional data in China *	
			Distribution of milk production	Distribution of milk processors
Location (*n* = 14)	Central-north	5 (36%)	30%	22%
	Northeast	3 (21%)	10%	13%
	Northwest	5 (36%)	12%	18%
	East	1 (7%)	24%	23%
	Others	0 (0%)	24%	25%
Number of employees (*n* = 11) **	100-500	6 (55%)		
	500-1000	2 (18%)		
	1000-10000	2 (18%)		
	> 10000	1 (9%)		
Age of company (*n* = 11) **	10-20	4 (36%)		
	20-30	4 (36%)		
	30-40	2 (18%)		
	> 40	1 (9%)		
Age of respondents (*n* = 9) **	20-30	1 (11%)		
	30-40	5 (56%)		
	40-50	2 (22%)		
	50-60	1 (11%)		
Working experience of the respondents (years) (*n* = 8) **	0-10	3 (38%)		
	11-20	3 (38%)		
	21-30	2 (25%)		
Annual production of liquid milk (ton) (*n* = 8) **	10,000-100,000	3 (38%)		
	100,000-1,000,000	4 (50%)		
	> 1,000,000	1 (12%)		

* Data retrieved from (Ministry of Agricultural P. R. China, 2016). ** Total participant number varies since some respondents did not provide the completed demographic information.

Table 3.3. The three key elements and 48 fraud factors of the food fraud vulnerability assessment.

Opportunities	Motivations	Controls
Technical opportunities	Economic drivers	Technical controls
1. Ease of adulteration	10. Supply and price of milk	30. Specificity and accuracy of the fraud monitoring system in place for raw milk own company (processors only)
2. Availability of technology for adulteration of the raw milk (processors only)	11. Valuable components and attributes	31. Systematics and autonomy of verification of the fraud monitoring system for raw milk own company (processors only)
3. Detectability of fraud in raw milk (processors only)	12. Economic health of the own company	32. Specificity and accuracy of the fraud monitoring system in place for the milk product in the own company
4. Availability of technology for adulteration of the milk product	13. Business strategy of the own company	33. Verification of fraud monitoring system for the milk product own company
5. Detectability of fraud in the milk product	17. Financial pressure imposed by the company on the farmers	34. Accuracy of the information system wrt mass balance control own company
	18. Supplier's (for the farmers, Customer's) economic health	35. Extensiveness of the tracing and tracking system in the own company
	19. Supplier's (for the farmers, Customer's) business strategy	39. Contractual requirements with supplier in the own company (for the farmers, customers)
	24. The economic health of the sector	40. Specificity and accuracy of the supplier's (for the farmers, customer) fraud monitoring system
	28. Level of competition in sector	41. Accuracy of the supplier's (for the farmers, customer) information system wrt mass balance control
	29. Price differences due to regulatory differences	42. Supplier's (for the farmers, customer) tracking and tracing system
		48. Availability of a fraud contingency plan
Opportunities in time and place	Cultural and behavioural drivers	Managerial controls
6. Accessibility to production activities	14. Ethical business own company	36. Application of integrity screening of employees in the own company
7. Transparency of the chain network	15. Previous irregularities of the own company	37. Strictness of the ethical code of conduct in the own company
8. Relationships within the supply chain	16. Corruption level of the country in which the own company is active	38. Support of a whistle blowing system in the own company
9. Historical evidence of milk fraud	20. Supplier's (for the farmers, Customer's) ethical business culture	43. Social control and transparency across the chain network
	21. Supplier's (for the farmers, Customer's) previous irregularities	44. Established guidance for fraud prevention and control in the sector
	22. Victimization of the supplier (for the farmers, Customer)	45. Specificity of the national food policy
	23. Corruption level of the country in which the supplier/customer is active	46. Strictness of law enforcement in the local chain
	25. Customer's previous irregularities (processors only)	47. Strictness of law enforcement in the international chain
	26. Sector ethical business culture	
	27. Historical evidence of milk fraud, within sector	

3

3.2.3. Data analysis

The questionnaire comprised 48 questions and 3 optional answers to each question. A three-level score system was used for the answers. These answers depicted typical descriptions and reflected low, medium and high vulnerability situations associated with the related fraud factors. For the factors related to opportunities and motivations, the answers with score 1, 2 and 3 reflected a low, medium and high vulnerability level, respectively. For the controls-related factors, the answers with the score 1, 2 and 3 reflected a low, medium and high level of adequacy of control measures, which related to a high, medium and low level of vulnerability, respectively. The percentages of low, medium and high vulnerability scores for each fraud factor and tier group were calculated. The answers of each tier were used to evaluate the perceived vulnerability of that tier only. The fraud factors 2-3, 25 and 30-31 were not applicable to the farmers, so they were left out for the assessment for the farmers.

To summarize the overall results, the scores of each tier group were balanced to allow them to contribute to the same extent: the weighted frequency of provided answers of each question (F_i) was determined by the following formula, Eq. (3.1)

$$F_i = \sum_j {x_{ij}}/{n_j} \qquad\qquad \text{Eq. (3.1)}$$

Where F_i is the frequency of score i (i = 1, 2, 3), x_{ij} is the number of observations which get score i in group j (j = farmers, processors), n_j is the total number of observations in group j. The score with the highest F_i for the common fraud factors were used generate the radar charts and present the overall results of the assessment.

As our study results in ordinal data, multiple correspondence analysis (MCA) was applied for exploratory analysis, and non-parametric Kruskal-Wallis tests were applied for the group comparisons and Mann-Whitney U-tests for pairwise comparisons ($p < 0.05$ was considered significant). MCA was performed by R 3.4.0 (R Foundation for Statistical Computing, Vienna, Austria), and the Kruskal-Wallis tests and Mann-Whitney U-tests were performed by SPSS v23.0 (IBM Statistics Inc., Armonk, NY, USA).

Only the common factors of both tier groups (n = 34) were used for making the radar chart, performing MCA and the statistical comparisons, i.e. the fraud factors 2-3, 18-22, 25, 30-31, 39-42 were left out since they were not comparable for the farmer and processor groups.

3.3. Results and discussion

3.3.1. Overall fraud vulnerability in the Chinese milk supply chain

The overall perceived vulnerability of the Chinese milk supply chain is presented by means of the modes for each fraud factor in Fig. 3.1. Furthermore, the relative frequencies of low, medium and high vulnerability scores for the assessments, overall and for each tier group, are shown in Fig. 3.2. The results are discussed below.
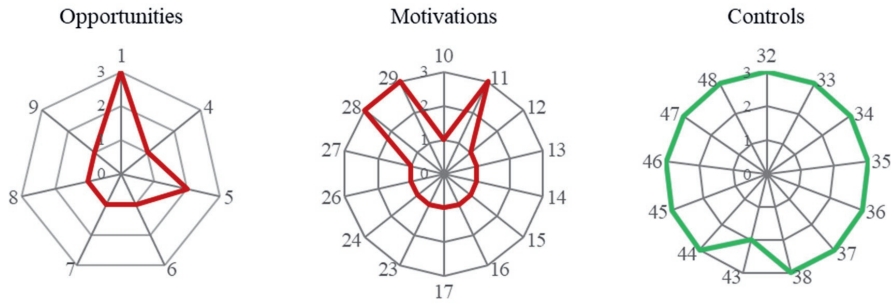
Fig. 3.1. Radar charts of modes for common fraud factors in the fraud vulnerability assessments over all respondents for opportunities (fraud factors 1-9), motivations (fraud factors 10-29), and control measures (fraud factors 32-48). Modes were weighted for the two tier groups to balance for group size. Explanation of fraud factor numbers is given in Table 3.3. Fraud factors which were not included in the assessments of both tier groups (factors 2-3, 18-22, 25, 30-31, 39-42) were omitted.
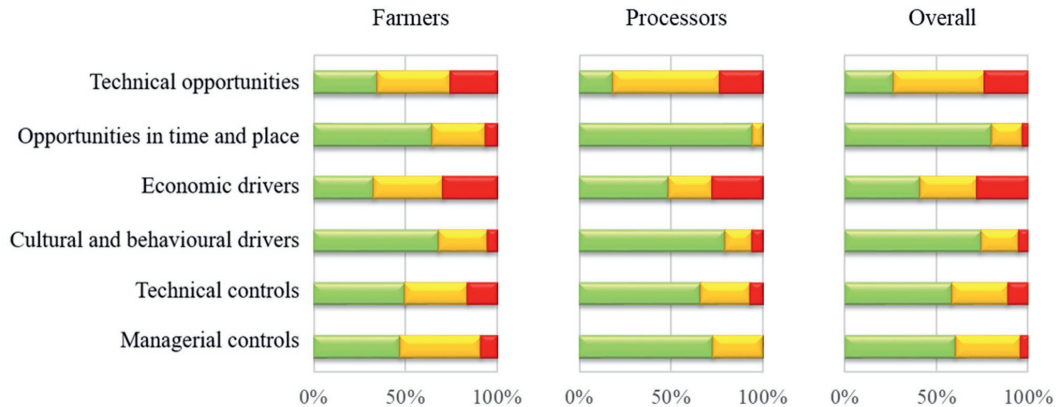
Fig. 3.2. The relative frequencies of vulnerability scores of fraud factors from all respondents for the six fraud factor categories (technical opportunities, opportunities in time and space, economic drivers, cultural and behavioural drivers, technical controls and managerial controls) and for the two tier groups (for farmers $n = 90$, for processors $n = 14$). Overall results present weighted frequencies for farmers and processors. The low, medium and high vulnerability portions are coloured green, orange and red, respectively.

Opportunities. The radar chart of opportunities-related factors (Fig. 3.1) shows that the ease of adulteration (fraud factor 1) and detectability of milk fraud (fraud factor 5) were rated as medium to high risk and thus were considered to contribute more to overall fraud vulnerability than the other factors, which were all rated low risk on average. It is widely acknowledged that milk, as a liquid material with complex composition, is easy to manipulate (Jack, 2015; Yang et al., 2019). There are various on-site methods available for fraud screening, such as rapid detection of melamine, aflatoxin, antibiotic and veterinary drug residue (Jaiswal, Jha, Kaur, Borah, & Ramya, 2018; Karczmarczyk, Baeumner, & Feller, 2017; McGrath et al., 2015; Naik et al., 2017; Zhou et al., 2014). However, the confirmation of milk authenticity still

needs advanced techniques in the laboratory and is, therefore, usually time-consuming. The potential fraudster might take advantage of the ease of adulteration together with the weakness of detectability to perpetrate irregularities out of sight. Overall, there seem to be some technical opportunities, but the fraud opportunities in time and place for milk production are relatively low (Fig. 3.2).

Motivations. Among the various motivations-related factors studied, the economic drivers were perceived as medium risk level, and cultural and behavioural drivers were perceived as low risk level on average (Fig. 3.2). The intensive competition across the sector has put high pressure on the participants to achieve their goals. This competition is both domestic and international. There are more than 600 dairy processors in China, and some of them are in a predominant position (Li, 2016). The economy scale and the marketing power allow the large size companies to pursue aggressive business strategies (Xiu & Klein, 2010). On the contrast, the medium and small processors face large challenges to survive in the fierce competition. In addition, the increasing international trade makes the competition more intense. The drop of trust in domestic milk after the melamine incident drove groups of consumers to prefer imported milk products (El Benni et al., 2019), which further intensified the competition across the sector nationally.

Pricing differences due to regional differences or global supply shortages can increase the fraud vulnerability (Moyer, De Vries, & Spink, 2017). Prices of raw milk and milk products in China have been higher than those of other countries for many years. This price gap may motivate milk processors to use cheaper milk powder as replacements, which in turn may increase the competition at farm stage and make the milk more vulnerable to fraud.

Controls. Most controls are considered as being highly adequate by the milk actors on average (Fig. 3.1). The well-established control measures can counteract the vulnerability generated from opportunities and motivations (van Ruth et al., 2017), and role-play as deterrence for the potential offenders. An appropriate deterrence mechanism can stimulate fraud mitigation move from fraud detection to fraud prevention (Manning, 2016). Our study shows that the social control and transparency of the network (fraud factor 43) is considered less adequate. Social control, which can be in various forms, such as promise keeping, information exchanging and cooperative problem solving, can, to some extent, substitute formal control in domestic buyer-supplier business relationships in China (Li, Xie, Teo, & Peng, 2010). Zhang et al. (2015) highlighted that the consumers' high tolerance of illegal behaviour and the lack of cross-market defence for dishonest companies in China would make Chinese food enterprises easier victims of food fraud. This is in line with our results.

To summarize and considering overall, average results, the milk supply chain in China is considered low-medium vulnerable to food fraud. However, there are differences between actors, tier groups and groups with different business characteristics. These matters will be further discussed in the next sections.

3.3.2. Differences in fraud vulnerability between tiers

An overview of the MCA for the fraud vulnerability of farmers and processors is presented in Fig. 3.3. We see that the farmers group is more widespread, whereas the processors group is more distinct and overlaps with some of the farmers. To examine statistical significance of the differences between the two tier groups, the mean rank for the two tier groups for each factor was present (Table 3.4). Bar charts depicting the relative frequencies of low, medium and high scores are available as Supplementary material (Fig. S3.2A-S3.2C).

Table 3.4. Mean ranks of scores of the fraud factors from the food fraud vulnerability assessments for the farmers and processors, and the statistical relevance of differences between the two tier groups. Fraud factors which were not included in the assessments of both tier groups (factors 2-3, 18-22, 25, 30-31, 39-42) were omitted. Explanation of fraud factor numbers is given in Table 3.3. High ranks for opportunities and motivations, and low ranks for controls indicate higher vulnerability.

Fraud element	Fraud factor no.	Farmers ($n = 90$)	Processors ($n = 14$)	p value
Opportunities	1	52	55	0.753
	4	51	64	0.076
	5	54	46	0.312
	6	53	50	0.548
	7	56	33	0.002 *
	8	56	28	< 0.001 *
	9	54	45	0.101
Motivations	10	55	38	0.034 *
	11	53	53	1.000
	12	56	29	< 0.001 *
	13	54	43	0.051
	14	55	38	0.011 *
	15	53	50	0.368
	16	55	40	0.043 *
	17	55	37	0.021 *
	23	54	44	0.160
	24	55	36	0.010 *
	26	54	42	0.108
	27	51	65	0.060
	28	50	68	0.027 *
	29	53	51	0.848
Controls	32	50	66	0.047 *
	33	52	56	0.599
	34	51	63	0.101
	35	50	68	0.020 *
	36	49	74	0.002 *
	37	49	75	0.001 *
	38	50	68	0.028 *
	43	51	62	0.135
	44	52	59	0.350
	45	50	68	0.022 *
	46	51	62	0.085
	47	52	54	0.800
	48	49	77	< 0.001 *

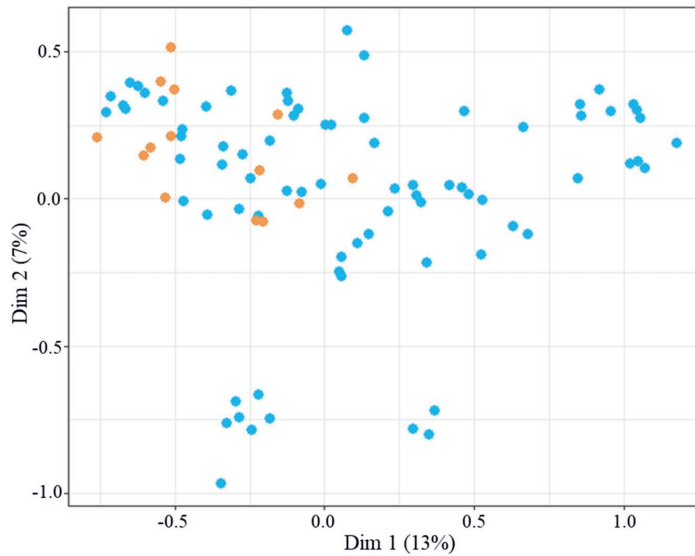* Significant difference in a row (Mann-Whitney U-test, $p < 0.05$).

3

Fig. 3.3. Scores plot of the first two dimensions of multiple correspondence analysis on the food fraud vulnerability assessment data of the farmers (blue symbols) and processors (orange symbols). Fraud factors which were not included in the assessments of both tier groups (fraud factors 2-3, 18-22, 25, 30-31, 39-42) were omitted.

Eighteen common factors show no significant differences (Mann-Whitney U-test, $p > 0.05$) between the two tier groups (Table 3.4). Some of these factors contribute highly to the fraud vulnerability. These factors have already been discussed in section 3.3.1. when the average vulnerability profile of the Chinese milk chain was discussed. On the other hand, there are 16 common factors that show significant differences (Mann-Whitney U-test, $p < 0.05$) between the perspectives of the farmers and processors. They include two opportunities-related factors, seven motivations-related factors and seven controls-related factors (Table 3.4).

Opportunities. The opportunities in time and place related factors transparency and relationship in the supply chain were rated significantly different by the farmers and processors. Most farmers indicated they knew their direct suppliers and customers well, but they had very limited information on the other actors in the supply chain. When the demand and supply became imbalanced about ten years ago, the farmers might have sold raw milk to short term partners or even independent milk brokers (Gale & Hu, 2009). The flexible business relationships require a high level of quality control and can increase fraud vulnerability in certain circumstances. On the contrary, the processors seemed to have more power to exchange information with other nodes in the chain, and have long term relationships with their business partners. One reason might be that a certain number of large-sized milk processors have set up the so-called "enterprise plus farmers" business model to secure milk supply and strengthen the farm management (NBSO, 2016). As a result, the processors can maintain long term relationships, and consequently could lower the risk caused by varying business relationships.

Motivations. Farmers perceived more (extensive) economic drivers and cultural and behavioural drivers than the processors (Fig. 3.2). Interestingly, these concern five economic driver related factors and two cultural and behavioural driver related factors (Table 3.4).

In regard to economic drivers, the two tiers differ particularly in opinion regarding the pricing of milk, economic health and financial pressure. All farmers presented significantly higher risk scores than the processors. This makes sense since the farm gate milk price is greatly influenced by various factors, such as supply and demand, milk composition and attributes, feed cost, labour cost and machine maintains, etc. (GAO, 2004). However, the most prevalent raw milk payment mechanism in China is that milk processors determine the price according to the enterprise standards and market supply and demand. The milk companies have more power in price bargaining in this mechanism (Xiu & Klein, 2010). The maximization of profit in this system drives the processors to set the farm gate milk price marginally over the cost of producing, which results in considerable financial pressure for the producers (farmers). Both the high costs and low raw milk price make it more difficult for the farmers to achieve their expected financial goals. This is in agreement with the fact that the power of production, processing and marketing in the Chinese milk industry is imbalanced between farmers and processors (Wu et al., 2018). In contrast with the farmers' perspective, the majority of the processors considered that their suppliers (the farmers) experienced no serious financial pressure from them, the reason of which was that the processors believed the price could cover the cost of raw milk production. Farmers and processors also perceived the level of competition differently. This may be due to the current transition in the Chinese farming structure. This structure is shifting from being decentralized and scattered to intensive and large-size breeding (Wu et al., 2018). This transition might cause fierce competition among certain farms but does not directly affect the processors.

Regarding the cultural and behavioural drivers, farmers and processors differed significantly in perceiving the factors ethical business culture of the own company and the corruption level of the country they are active in. The farmers rated higher risk for both. Some of the farmers are smaller sized businesses than the processors. Small owner-managed businesses are very different in organisation compared to large corporations where ownership and management are separated. Although studies have not shown that these differences result in a direct different degree of (un)ethical behaviour, it has to be considered that the ethical values and inclinations of the small business owner will have far more direct consequences on the practices of the business as a whole (Longenecker, Moore, Petty, Palich, & McKinney, 2006).The difference in scores for the corruption level factor is more surprising, since all interviewed actors are active in China, and they may refer to the same corruption perception index. The gap between the two tier groups in rating the corruption level is probably due to the fact that farmers need to deal with different governmental agencies than processors, therefore they have different experiences with corruption. The difference in perceiving the corruption level may reflect the situated acceptance of corruption as part of doing business by the two tiers. It has been pointed out that the acceptability of (non-)criminal behaviour is one of the criteria for determining risk of food crime (Manning, Smith, & Soon, 2016). Besides that, the

respondents perceived China less corrupted than the result from Transparency International (2018) report. This might be explained by the desirability bias, that in social research regarding sensitive topics, providing three optional answers increases the respondents preference to project a favourable image to others and avoid embarrassment (Fisher, 1993).

To summarize, all economic and cultural and behavioural drivers that differed significantly in scores between farmers and processors, were rated higher by the farmers.

Controls. The scores of the farmers and processors differed significantly for six control factors, i.e. two technical controls and five managerial controls. For all more adequate controls were available perceived by the processors compared to the farmers.

First of all, the adequacy of the fraud monitoring systems and track and trace systems differed between the two tier groups. According to the interviews with the farmer experts, the farmers might easily consider that the fraud monitoring system was designed to meet the requirement of the food safety management system. Consequently, most farmers perceived the fraud monitoring system as medium-high adequacy. For instance, these measures require the raw milk to be recalled if it were found to be adulterated or contaminated. Most farmers perceived that they set up integrity screening for employees at key positions, established general ethical codes of conduct and had simple whistle blowing systems. However, it should be emphasized that over 80% of the farms in this study had more than ten employees, and the lack of systematic managerial controls in such intensive farming businesses adds to fraud vulnerability. For instance, scholars have stressed that an appropriate whistleblowing strategy is important to safeguard individuals, mitigate food fraud, and protect consumers from potential harm (Soon & Manning, 2017). The processors perceived themselves having more adequate controls in place than the farmers (Table 3.4), as may be expected from businesses in the middle of the chain that may be victimized more easily by others.

Surprisingly, the coverage of food fraud by the national policy to cover was perceived differently by the two tier groups. The processors believed it was highly adequate, whereas the farmers believed it covered food fraud only generally. The melamine incident has raised attention to cover food fraud in the regulatory framework. Therefore, series of regulations and policies have been released by several national departments (Supplementary material, Table S3.1) for the milk processing enterprises. These regulations emphasise quality assurance for the processors, including record keeping for mass balance control, setting up a complete track and trace system, etc. However, it is recognized that some of the food policies aiming at food fraud are not enforced at farm level but only at the processor level explaining the different perceptions of the two tier groups. Thus, altogether, less adequate technical and managerial controls make the farmers more susceptible to food fraud than the milk processors.

Summarizing section 3.3.2, it is obvious that the farmers are more vulnerable to fraud than the processors for a wide range of indicators. The enhanced vulnerability is due to increased opportunities, increased motivations and less sufficient control measures.

3.3.3. Differences in fraud vulnerability between farms

The large number of farms allowed a further investigation into the relationship between farm business characteristics and fraud vulnerability. For a first exploration, MCA was carried out on the farm data to explore similarities and differences in the assessment score profiles. MCA plots, presented in Fig. 3.4, illustrate the effect of geographical location and business size. The plots show some clusters; for instance, the north-western farms (purple symbols) and a group of small sized central-north farms (green round symbols) show a distinct pattern. In order to research the influence of location and business size further, differences in factor scores between groups with different characteristics were examined for their significance (Tables 3.5 and 3.6).

Table 3.5. Mean ranks of scores of the fraud factors from the food fraud vulnerability assessments for the farmers, and the statistical relevance of differences between the farms based in different locations [a]. Explanation of fraud factor numbers is given in Table 3.3. High ranks for opportunities and motivations, and low ranks for controls indicate higher vulnerability.

Fraud element	Fraud factor no.	N ($n = 54$)	E ($n = 16$)	NE ($n = 10$)	NW ($n = 10$)	p value
Opportunities	4	40 b	42 ab	60 a	64 a	0.002 *
	5	49 a	43 ab	49 ab	27 b	0.031 *
	6	42 b	46 ab	58 a	50 ab	0.019 *
	8	53 a	39 ab	28 b	34 ab	0.001 *
Motivations	10	42 b	41 b	45 ab	70 a	0.007 *
	11	54 a	34 b	34 b	29 b	< 0.001 *
	17	52 a	33 b	42 ab	34 ab	0.009 *
	18	41 b	51 ab	43 ab	65 a	0.010 *
	22	47 a	49 a	25 b	55 a	0.018 *
	24	42 b	44 ab	65 a	48 ab	0.040 *
	28	35 c	65 a	47 bc	66 ab	< 0.001 *
	29	40 b	56 ab	40 ab	62 a	0.005 *
Controls	36	40 b	48 ab	53 ab	65 a	0.015 *
	37	42 b	40 ab	52 ab	66 a	0.020 *
	38	40 b	55 ab	39 b	68 a	0.003 *
	45	42 b	62 a	37 b	48 ab	0.008 *

[a] N, E, NE, and NW stands for central-north, east, northeast and northwest of China, respectively. * Different letters behind ranks in a row indicate significant different (Kruskal-Wallis test followed by Mann Whitney U test, $p < 0.05$).

Table 3.6. Mean ranks of scores of the fraud factors from the food fraud vulnerability assessments for the farmers, and the statistical relevance of differences between the farms of different sizes [a]. Explanation of fraud factor numbers is given in Table 3.3. High ranks for opportunities and motivations, and low ranks for controls indicate higher vulnerability.

Fraud element	Fraud factor no.	S ($n = 26$)	M ($n = 32$)	L ($n = 32$)	p value
Opportunity	1	58 a	39 b	42 b	0.009 *
Motivations	12	36 b	51 a	48 ab	0.036 *
	14	56 a	40 b	43 ab	0.015 *
	28	33 b	46 a	55 a	0.002 *
	29	35 b	54 a	46 ab	0.005 *
Controls	35	44 ab	38 b	54 a	0.025 *
	38	36 b	46 ab	52 a	0.045 *
	42	35 b	52 a	47 ab	0.015 *
	45	36 b	45 ab	54 a	0.013 *
	48	30 b	46 a	57 a	< 0.001 *

[a] S, M, and L stands for small, medium and large sized farm, respectively. * Different letters behind ranks in a row indicate significant differences (Kruskal-Wallis test followed by Mann-Whitney U-test, $p < 0.05$).
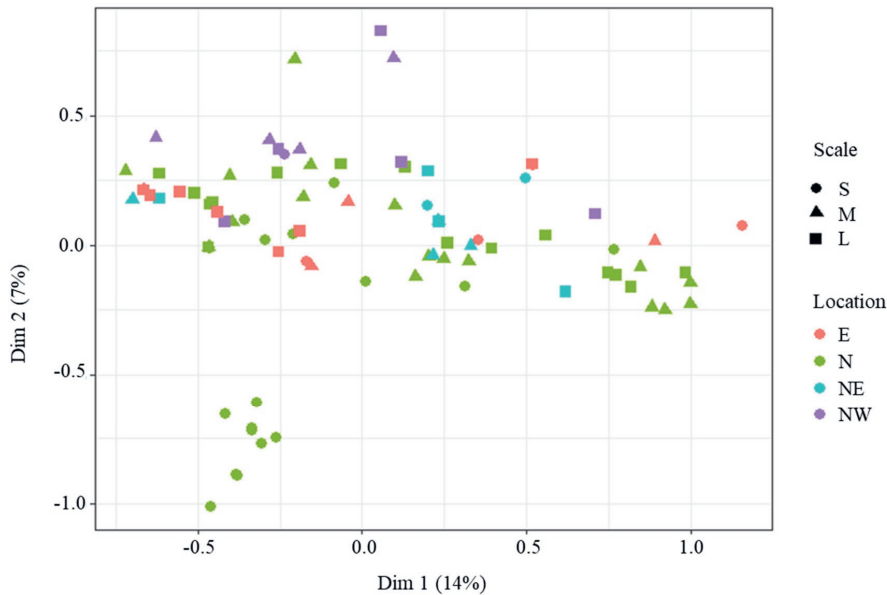
Fig. 3.4. Scores plot of the first two dimensions of multiple correspondence analysis on the food fraud vulnerability assessment data of the farmers coded according to their business characteristics. S, M, and L stand for small, medium, and large size enterprises, respectively. E, N, NE, and NW stand for east, central-north, northeast, and northwest of China, respectively.

3.3.3.1. Farm geography

Opportunities. Although four opportunities-related factors show significant differences between farmers from the different geographical areas (Table 3.5), the mutual differences appear of a more diverse nature. It seems that the farmers from the central-north considered the

complexity of detectability of adulteration (fraud factor 5) as a higher risk factor as well as the relationships within the supply chain (fraud factor 8). Whereas those from the northeast and northwest rated the available technology of adulteration (fraud factor 1) and accessibility of farming activities (fraud factor 6) higher than their counterparts.

Motivations. The results on motivations-related factors reveal that the farmers from the northwest considered the supply and price of milk to be very fluctuating, the competition as very intensive and they felt more financial pressure by their customers (fraud factors 10, 17, 28) than other farmers. Consequently, it makes sense that they were also worried about their economic health (fraud factors 18, 24). The current pricing system of raw milk, which is determined by processors, might be an underlying cause of the price fluctuation in northwest China.

Controls. It appears that the farms from the northwest of China had more extensive managerial controls (fraud factors 36-38) in place than the farmers from the central-north of the country. Furthermore, the farmers from the east were more satisfied with the national food policy, whereas those from the central-north and northeast of the country believed the policy could target fraud more specifically.

Geographical location of the farms seems to affect fraud vulnerability resulting from a number of fraud factors. However, not in one direction. No literatures were found in term of ranking the fraud vulnerabilities for the farmers from different geographical locations. Thus, one cannot pinpoint one extremely vulnerable region.

3.3.3.2. Farm size

Opportunities and motivations. The respondents from small sized farms considered milk adulteration to be more complex than those from medium and large farms (fraud factor 1). This may be related to general knowledge level available at the farms. It is interesting that the small sized farmers appeared in better economic health (fraud factor 12) and had to deal with a lower level of competition (fraud factor 28) than the medium sized farmers (Table 3.6). This is in agreement with data available on profitability of different sizes of farms. Studies showed that the average return on capital of small sized farms is higher (31%) than that of the medium sized ones (27%), whereas the return rate of large size farms is somewhere in between (Ministry of Agriculture P. R. China, 2016). Diverse costs including expense of feed, labour, land, etc. could impact the profitability of the farm (Wang, Liu, Makkar, Wei, & Xu, 2014) and may differ with farm size. As a result, it appears that fraud vulnerability resulting from economic drivers is perceived higher by the medium sized farms.

Controls. The controls related to the availability of a track and trace system (fraud factor 35), a whistle blowing system (fraud factor 38) and a contingency plan (fraud factor 48) were more adequate in the large and medium sized farms than at small farm level (Table 3.6). Since ca. 70% of the small farms are family businesses, it is likely that less budget for such control measures is available. This is in agreement with another survey which showed that the small

sized farms in the northern area of China have less developed practices regarding information recording (Yang, Chen, & Kong, 2019). Small farms also had the impression that the national food policy is not covering food fraud well, whereas this was less pronounced for other sized farms. Small farms may not be in touch with authorities as often as the larger ones. To summarize, less adequate controls result in a higher fraud vulnerability of small sized farms.

3.4. Concluding remarks

The milk supply chain in China is, on average, perceived as low to medium vulnerable to food fraud according to the results of the current study. However, considerable differences between actors' perspectives exist. Farmers are more vulnerable than processors because of vulnerabilities resulting from increased opportunities, increased motivations and implemented less adequate control measures. When zooming in on the farmers, it appears that both farm location and size affect the perceived vulnerability profiles. Farmers in the northwest present for instance more economic drivers, and smaller sized farms have fewer controls.

Compared to the situation of the milk chain in the Netherlands (Yang et al., 2019), the overall perceived fraud vulnerability seems fairly similar in China and is for instance lower than those of the spice or olive oil supply chain networks (van Ruth et al., 2017). There are differences between the perspectives of the tier groups in the two countries, however. The processors consider themselves more vulnerable in the Netherlands, while the farmers take on this role in China. Obviously, threats can come from outside or from within organizations (van Ruth et al., 2017). For farmers this would be more likely to come from within the organisation, whereas processors can be offenders themselves but can also be victimized. A more detailed comparison is difficult because of the major differences in traditional culture, political system, etc. between China and the Netherlands and need to be taken into account as well.

3.5. Considerations and recommendations

The SSAFE FFVA tool is originally designed as a self-assessment to determine the vulnerabilities of an individual business. It is easy to use by the external examiners to get an overall picture. However, for the research of in-depth analysis of sensitive factors, for instance ethical business culture and corruption level, the social desirability bias needs to be considered. The respondents are more likely to choose the answer to be viewed favourably by the others. For the research on further evaluation of the sensitive factors, it is recommended to use indirect questions or a Likert scale to reduce this bias.

In addition, this tool is developed in the western context. Chinese business culture, which emerged from Confucianism impacted by Socialism, emphasise on the relation of the family, the community and the greater society differs from Western traditions (Ip, 2009; Wu & Davidson, 2011; Zhang, Cone, Everett, & Elkin, 2011). It is recommended to consider the difference between the Chinese and Western economic and cultural practices to further evaluate cultural factors such as ethical business culture.

The current study provides an insight on the perceived fraud vulnerability and related fraud factors of dairy farmers and processors in China. In future work, it is of interest to carry out a wider food fraud vulnerability assessment for the other actors in the milk production chain, including e.g. feed suppliers, milk collectors, milk derived product manufacturers, retailers, food service, etc. to consider the full breadth of the Chinese milk supply chain.

Acknowledgements

3

References

Cadieux, B., Goodridge, L. D., & Spink, J. (2019). Gap analysis of the Canadian food fraud regulatory oversight and recommendations for improvement. *Food Control, 102*, 46-55.

El Benni, N., Stolz, H., Home, R., Kendall, H., Kuznesof, S., Clark, B., Dean, M., Brereton, P., Frewer, L. J., Chan, M., Zhong, Q., & Stolze, M. (2019). Product attributes and consumer attitudes affecting the preferences for infant milk formula in China – a latent class approach. *Food Quality and Preference, 71*, 25-33.

Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research, 20*(2), 303-315.

Gale, H. F. Jr., & Hu, D. (2009). Supply chain issues in China's milk adulteration incident. Paper presented at *the International Association of Agricultural Economists' 2009 Conference*, August 16-22, 2009. Beijing, China. Retrieved from http://purl.umn.edu/51613.

GAO. (2004). Dairy industry: Information on milk prices, factors affecting prices, and dairy policy options. Retrieved from http://www.gao.gov/products/GAO-05-50.

Graham-Harrison, E. (2009). China's tainted milk parents hope for death penalty. Retrieved from https://www.reuters.com/article/idUSPEK294403.

Ip, P. K. (2009). The challenge of developing a business ethics in China. *Journal of Business Ethics, 88*(1), 211-224.

Jack, L. (2015). *Risk Modelling of Food Fraud Motivation: 'NSF Fraud Protection Model' Intelligent Risk Model Scoping Project FS 246004: Final Report*. London: Food Standard Agency.

Jaiswal, P., Jha, S. N., Kaur, J., Borah, A., & Ramya, H. G. (2018). Detection of aflatoxin M1 in milk using spectroscopy and multivariate analyses. *Food Chemistry, 238*, 209-214.

Karczmarczyk, A., Baeumner, A. J., & Feller, K. (2017). Rapid and sensitive inhibition-based assay for the electrochemical detection of Ochratoxin A and Aflatoxin M1 in red wine and milk. *Electrochimica Acta, 243*, 82-89.

Lam, H., Remais, J., Fung, M., Xu, L., & Sun, S. (2013). Food supply and food safety issues in China. *The Lancet, 381*(9882), 2044-2053.

Levitt, S. D. & Dubner, S. J. (2014). *Think Like a Freak*. New York, NY: William Morrow.

Li, S. (2016). White paper on China dairy 2016. Retrieved from http://www.sdddc.org/en/download/detail-251.aspx.

Li, Y., Xie, E., Teo, H., & Peng, M. W. (2010). Formal control and social control in domestic and international buyer-supplier relationships. *Journal of Operations Management, 28*(4), 333-344.

Longenecker, J. G., Moore, C. W., Petty, J. W., Palich, L. E., & McKinney, J. A. (2006). Ethical attitudes in small businesses and large corporations: Theory and empirical findings from a tracking study spanning three decades. *Journal of Small Business Management, 44*(2), 167-183.

Manning, L. (2016). Food fraud: Policy and food chain. *Current Opinion in Food Science, 10*, 16-21.

Manning, L., Smith R., & Soon J. M. (2016). Developing an organizational typology of criminals in the meat supply chain, *Food Policy*, *59,* 44-54.

McGrath, T. F., McClintock, L., Dunn, J. S., Husar, G. M., Lochhead, M. J., Sarver, R. W.,Klein, F. E., Rice, J. A., Campbell, K., & Elliott, C. T. (2015). Development of a rapid multiplexed assay for the direct screening of antimicrobial residues in raw milk. *Analytical and Bioanalytical Chemistry, 407*(15), 4459-4472.

Ministry of Agriculture P. R. China (2016). *China Dairy Yearbook*. Beijing, China: China Agriculture Press.

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.

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.

Naik, L., Sharma, R., Mann, B., Lata, K., Rajput, Y. S., & Surendra Nath, B. (2017). Rapid screening test for detection of oxytetracycline residues in milk using lateral flow assay. *Food Chemistry, 219*, 85-92.

NBSO Jinan (2016). *Opportunity Report-Dairy Industry in Shandong Province*. Jinan, China. Retrieved from https://edepot.wur.nl/375801.

Ouyang, H. (2011). Five top safety issues of Chinese in 2011. *Insight China, 7*, 50-54.

Pei, X., Tandon, A., Alldrick, A., Giorgi, L., Huang, W., & Yang, R. (2011). The China melamine milk scandal and its implications for food safety regulation. *Food Policy, 36*(3), 412-420.

Qian, G., Guo, X., Guo, J., & Wu, J. (2011). China's dairy crisis: Impacts, causes and policy implications for a sustainable dairy industry. *International Journal of Sustainable Development & World Ecology, 18*(5), 434-441.

Soon, J. M., & Manning, L. (2017). Whistleblowing as a countermeasure strategy against food crime. *British Food Journal, 119*(12), 2630-2652.

Spink, J., & Moyer, D. C. (2011). Defining the public health threat of food fraud. *Journal of Food Science, 76*(9), R157-R163.

3

Spink, J. W. (2019). Criminology theory (Part 1 of 2): Foundational concepts. In *Food Fraud Prevention: Introduction, Implementation, and Management*. New York, NY: Springer.

SSAFE. (2017). Food fraud vulnerability assessment tool. Retrieved from http://www.ssafe-food.org/our-projects/.

Transparency International (2018). Corruption perceptions index 2018. Retrieved from https://www.transparency.org/cpi2018.

van Ruth, S. M., Huisman, W., & Luning, P. A. (2017). Food fraud vulnerability and its key factors. *Trends in Food Science & Technology, 67*, 70-75.

van Ruth, S. M., Luning, P. A., Silvis, I. C. J., Yang, Y., & Huisman, W. (2018). Differences in fraud vulnerability in various food supply chains and their tiers. *Food Control, 84*, 375-381.

Wang, C., Liu, J., Makkar, H. P. S., Wei, N., & Xu, Q. (2014). Production level, feed conversion efficiency, and nitrogen use efficiency of dairy production systems in China. *Tropical Animal Health and Production, 46*(4), 669-673.

Wang. D. (2009). Dairy industry loss 20 billion because of melamine scandal: Mengniu and Yili face heavier pressures. *South Metropolitan Newspaper*. Retrieved from http://news.hexun.com/2008-10-31/110843063.html.

Wu, J., & Kirk Davidson, D. (2011). The business-government-society relationship: A comparison between China and the US. *Journal of Management Development, 30*(1), 112-125.

Wu, X., Lu, Y., Xu, H., Lv, M., Hu, D., He, Z., Liu, L., Wang, Z., & Feng, Y. (2018). Challenges to improve the safety of dairy products in China. *Trends in Food Science & Technology, 76*, 6-14.

Xiu, C., & Klein, K. K. (2010). Melamine in milk products in China: Examining the factors that led to deliberate use of the contaminant. *Food Policy, 35*(5), 463-470.

Yang, X., Chen, K. Z., & Kong, X. (2019). Factors affecting the adoption of on-farm milk safety measures in northern China: An examination from the perspective of farm size and production type. *Journal of Integrative Agriculture, 18*(2), 471-481.

Yang, Y., Huisman, W., Hettinga, K. A., Liu, N., Heck, J., Schrijver, G. H., Gaiardoni, L., & van Ruth, S. M. (2019). Fraud vulnerability in the Dutch milk supply chain: Assessments of farmers, processors and retailers. *Food Control, 95*, 308-317.

Zhang, H., Cone, M. H., Everett, A. M., & Elkin, G. (2011). Aesthetic leadership in Chinese business: A philosophical perspective. *Journal of Business Ethics, 101*(3), 475-491.

Zhang, M., Qiao, H., Wang, X., Pu, M., Yu, Z., & Zheng, F. (2015). The third-party regulation on food safety in China: A review. *Journal of Integrative Agriculture, 14*(11), 2176-2188.

Zhang, W., & Xue, J. (2016). Economically motivated food fraud and adulteration in China: An analysis based on 1553 media reports. *Food Control, 67*, 192-198.

Zhou, C., Zhang, X., Huang, X., Guo, X., Cai, Q., & Zhu, S. (2014). Rapid detection of chloramphenicol residues in aquatic products using colloidal gold immunochromatographic assay. *Sensors, 14*(11), 21872-21888.

Zhu, X., Huang, I. Y., & Manning, L. (2019). The role of media reporting in food safety governance in China: A dairy case study. *Food Control, 96*, 165-179.

3

Supplementary material

Table S3.1. List of Chinese regulations and policies related to the dairy production after 2008.

Category	Item	Issued by	Reference
Regulations for raw milk safety	Dairy quality and safety supervision and management regulations	State Council	(the Central Peoples's Government of the P. R. China, 2008)
	Administrative measures for the fresh milk procurement	Ministry of agriculture	(Ministry of Agriculture P. R. China, 2008)
	Notice Relating to Further Strengthening the Safety of Dairy Product Quality	General Office of the State Council	(the Central Government of Poeple's Republic of China, 2010)
Regulations for milk processing	Provisions about Heavier Punishment on Illegal Acts during the Dairy Cattle Breeding, Procurement and Transportation of Fresh Milk	Ministry of agriculture, National Development and Reform Commission	(Ministry of Agriculture P. R. China, 2011)
	Dairy Industry Rectifying and Revitalization Plan	State Council	(the Central People's Government of the P. R. China, 2008)
	Dairy Industry and industrialization Policies	National Development and Reform Commission	(National Development and Reform Commission, 2008)
	Food Safety Tractability Information Recording Criterion for Infant Formula Milk Powder Production Enterprises	State Food and Drug Administration	(State Food and Drug Administration, 2015)

References

Ministry of Agriculture P.R.China. (2008). Administrative measures for the fresh milk procurement. Retrieved from http://jiuban.moa.gov.cn/zwllm/tzgg/bl/200811/t20081121_1177078.htm.

Ministry of Agriculture P.R.China. (2011). Provisions about Heavier Punishment on Illegal Acts during the Dairy Cattle Breeding, Procurement and Transportation of Fresh Milk. Retrieved from http://www.moa.gov.cn/nybgb/2011/dwuq/201805/t20180521_6142675.htm.

National Development and Reform Commission. (2008). Dairy Industry and industrialization Policies. Retrieved from http://www.miit.gov.cn/n1146285/n1146352/n3054355/n3057292/n3057299/c3658081/content.html.

State Food and Drug Administration. (2015). Food Safety Tractability Information Recording Criterion for Infant Formula Milk Powder Production Enterprises. Retrieved from http://www.zyczs.gov.cn/html/bmgz/2018/9/1536722657155.html.

The Central Government of Poeple's Republic of China. (2010). Notice Relating to Further Strengthening the Safety of Dairy Product Quality. Retrieved from http://www.gov.cn/xxgk/pub/govpublic/mrlm/201009/t20100925_56777.html.

The Central People's Government of the People's Republic of China. (2008). Dairy Industry Rectifying and Revitalization Plan. Retrieved from http://www.gov.cn/jrzg/2008-11/19/content_1154518.htm.

The Central Peoples's Government of the People's Republic of China. (2008). Dairy quality and safety supervision and management regulations. Retrieved from http://www.gov.cn/zwgk/2008-10/10/content_1116657.htm**.**

Technical opportunities

Opportunities in time and place
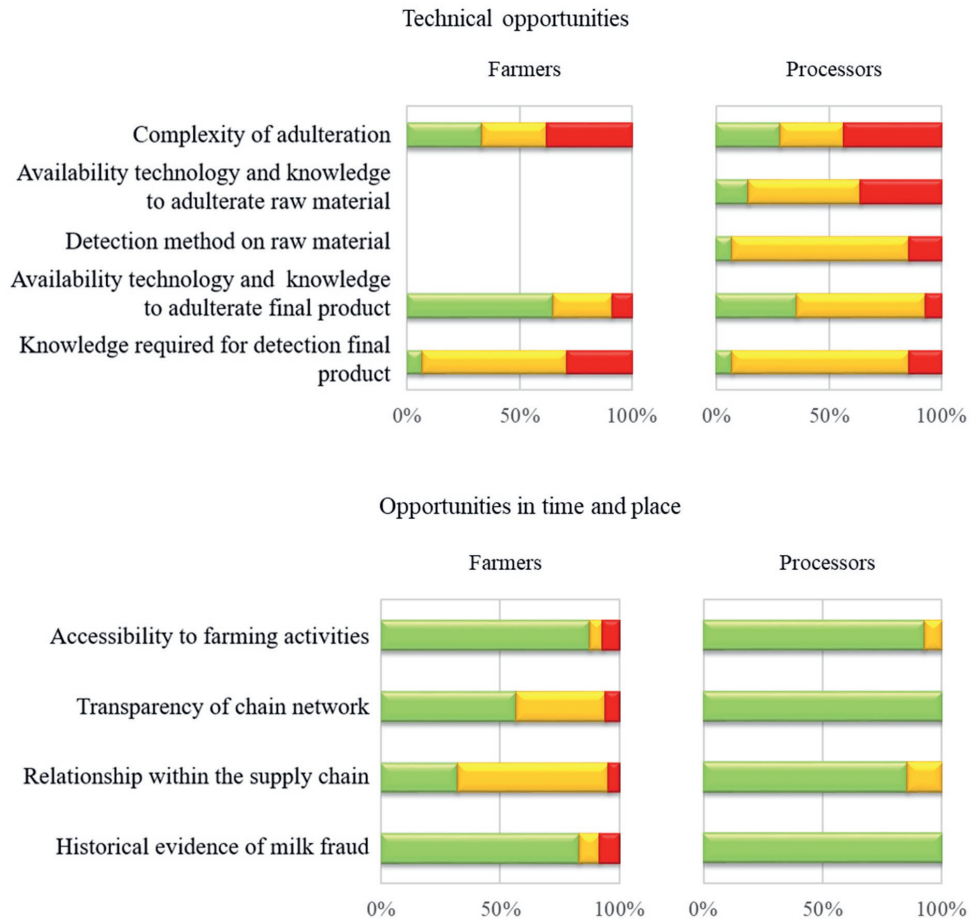
Fig. S3.2A. The relative frequencies of low, medium and high vulnerability scores for the opportunity related factors of the food fraud vulnerability assessment for the tier groups (for farmers $n = 90$, for processors $n = 14$). The low, medium and high vulnerability portions are coloured green, orange and red, respectively. The non-applicable factors are left blank.

Economic drivers

Cultural and behavioural drivers
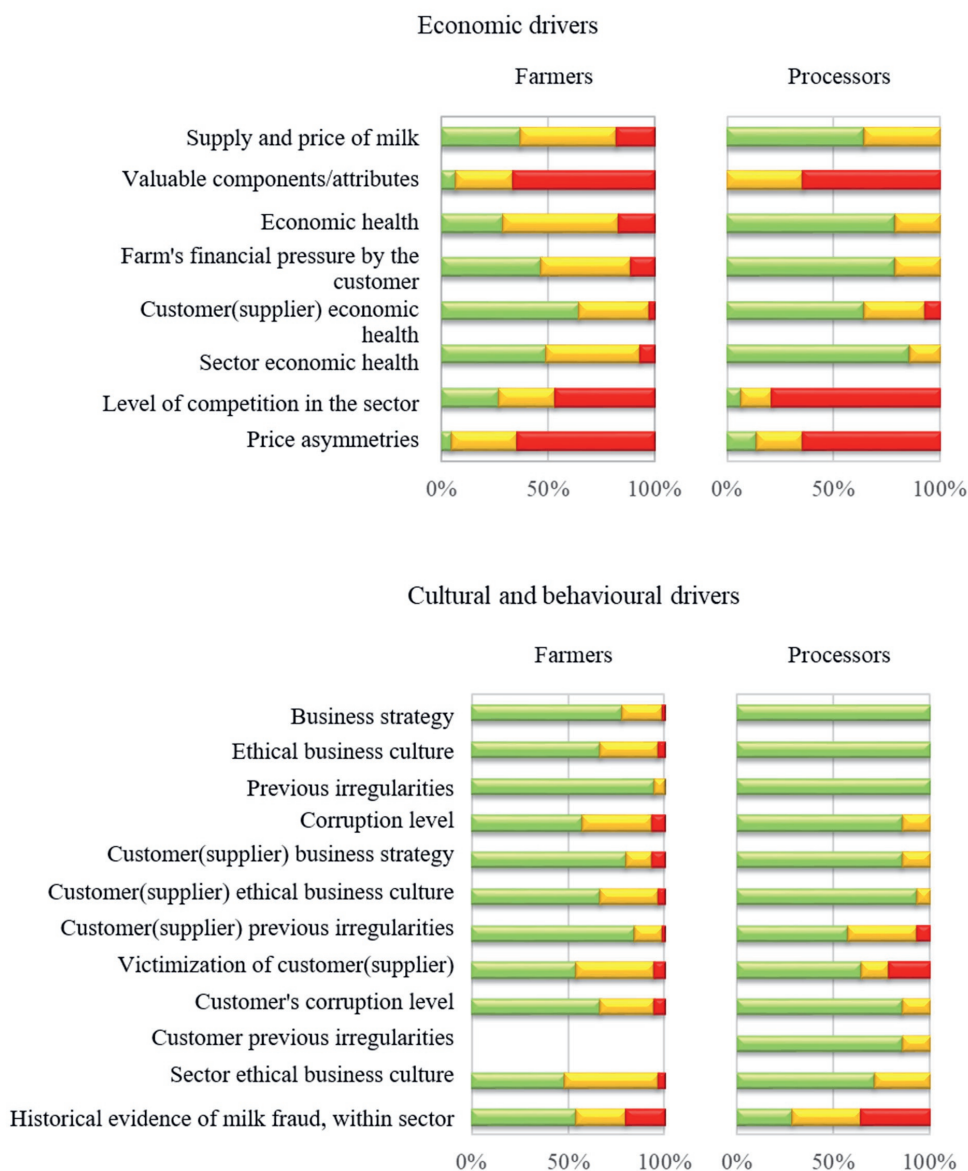
Fig. S3.2B. The relative frequencies of low, medium and high vulnerability scores for motivation related factors of the food fraud vulnerability assessment for the tier groups (for farmers $n = 90$, for processors $n = 14$). The low, medium and high vulnerability portions are coloured green, orange and red, respectively. The non-applicable factor is left blank.
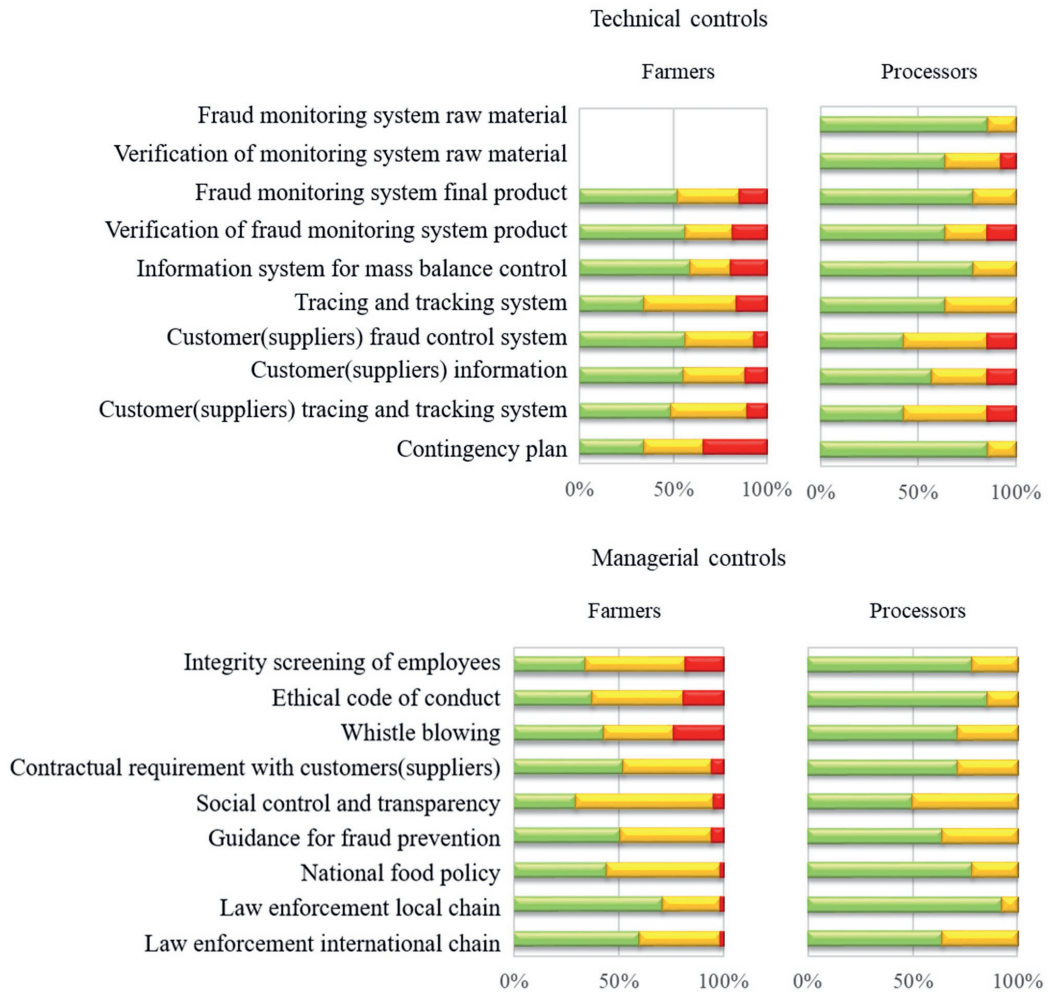
Fig. S3.2C. The relative frequencies of low, medium and high vulnerability scores for the control related factors of the food fraud vulnerability assessment for the tier groups (for farmers $n = 90$, for processors $n = 14$). The low, medium and high vulnerability portions are coloured green, orange and red, respectively. The non-applicable factors are left blank.

CHAPTER 4

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

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.

Keywords

Fourier transform infrared; Milk adulteration; Milk composition; Milkoscan measurements; One class classification; Profitability.

Chemical compounds cited 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).

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.

4

4.1. 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 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 FrieslandCampina, 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.

4.2. Materials and methods

4.2.1. 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.

4.2.2. 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 S4.1 (Supplementary material).

4.2.3. Adulterations

4.2.3.1. Single adulterations

The three milk pools prepared as described in section 4.2.1 were spiked with the adulterants according to Eq. (4.1) to Eq. (4.3). The weight of the protein-rich adulterants added to the 100 g milk pools was calculated according to Eq. (4.1):

$$Weight_{protein-rich\ adulterant} = 100\ g \times \left(\frac{Protein_{control} \times a\%}{Protein_{adulterant}} \right) \qquad \text{Eq. (4.1)}$$

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. (4.2):

$$Weight_{nitrogen\ adulterant} = 100\ g \times \left(\frac{Protein_{control} \times a\%}{f \times N_{adulterant}} \right) \qquad \text{Eq. (4.2)}$$

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. (4.3):

$$Weight_{carbohydrate\ adulterant} = 100\ g \times \left(\frac{TS_{control} \times a\%}{TS_{adulterant}}\right)$$ Eq. (4.3)

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 4.1. 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 minutes after the adulterants were added. In total, 294 single-spiked milk samples were thus prepared.

4.2.3.2. 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 4.2.2) 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 4.1. 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.

4.2.4. Measurements

All the samples were measured in duplicate using the MilkoScan FT120 instrument (Foss Electric, Hilleroed, Denmark), with wavenumbers of the spectrum from 5000-930 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.

Table 4.1. The amount of adulterants added into the three milk pools (per 100 g) for the single adulteration and combined adulteration [a].

Adulterant category	Adulterant	Single adulteration				Combined adulteration	
		Level 1 (g)	Level 2 (g)	Level 3 (g)	Level 4 (g)	Adulterant (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.

[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.

4.2.5. Statistical analysis

4.2.5.1. 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.

4.2.5.2. 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 0.5^{th} and 99.5^{th} 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. (4.4):

$$X_{adjusted} = \left(\frac{X - \mu}{\sigma} \times 2\sigma \right) + \mu \qquad\qquad \text{Eq. (4.4)}$$

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 0.5^{th} percentile and 99.5^{th} 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 S4.2 and S4.3 (Supplementary material).

4.2.5.3. 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 &

4

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. (4.5),

$$k\,(X1, X2)\;=\;Exp(-\,\gamma||X1 - X2||^2)$$ Eq. (4.5)

Accordingly, a SVM model performance depends on the model parameter γ.

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. (4.6):

$$Balanced\;accuracy = \frac{(True\;positive\;rate + True\;negative\;rate)}{2}$$ Eq. (4.6)

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).

4.2.6. 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. (4.7):

$$Net\;profit\;change\;=\;Increased\;profit\;-\;Cost\;of\;adulterant$$ Eq. (4.7)

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. (4.8):

$$Increased\ profit = \ 1000 \ \times \ \sum_i \Delta_i \times p_i \qquad\qquad \text{Eq. (4.8)}$$

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.

4.3. Result and discussion

4.3.1. Natural variation of the control samples

Variation in composition of the control milk samples was observed (Table 4.2; Supplementary material Table S4.2). 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, 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 4.2.5.2. Both datasets were used to evaluate the potential to detect suspected milk adulterations.

4.3.2. The impact of adulterants on the payment parameters

The impact of each adulterant on the apparent milk composition is presented in Table 4.3. 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. 4.1). 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.

Table 4.2. 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 [a].

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

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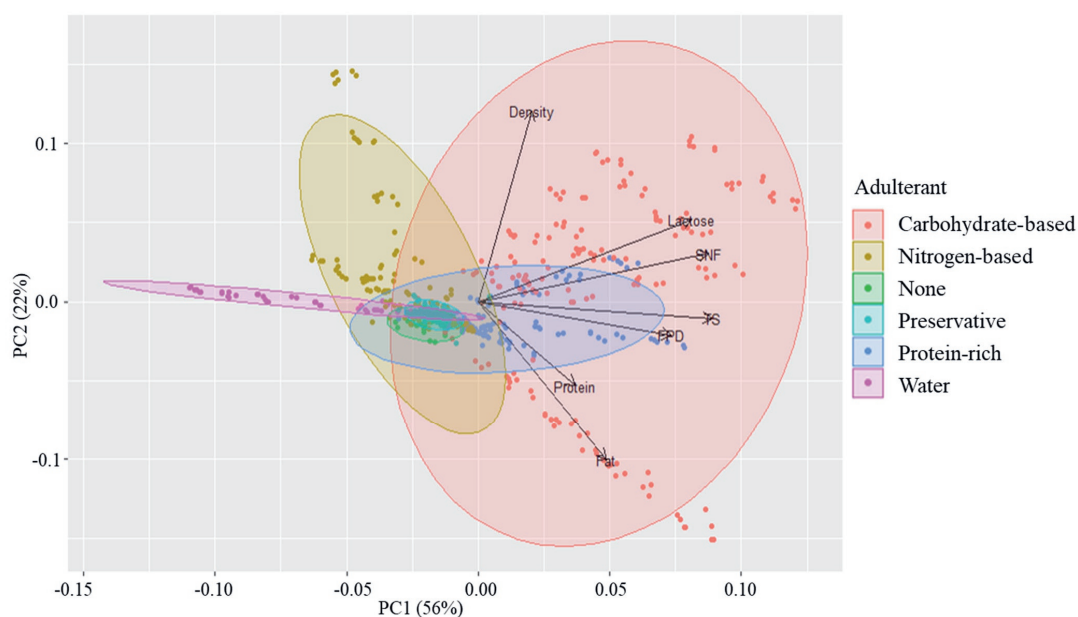

Fig. 4.1. The Principal Component Analysis (PCA) scores bi-plot of the first two PC dimensions of all milk samples based on the data obtained from the MilkoScan measurements.

Table 4.3. 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 S4.4 (Supplementary material) [a].

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

[a] The slope (*m*) and R-squared (R^2) 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.

4.3.2.1. 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 4.3, 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 4.3). 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 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 4.3 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 S4.1, 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.

4.3.2.2. 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 4.3 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.

4.3.2.3. 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 4.3, 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.

4.3.2.4. Impact of the preservatives addition

In the PCA plot (Fig. 4.1), 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 4.3).

4

4.3.3. When will the adulterations be detected?

4.3.3.1. 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. 4.2). 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. 4.2a). 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.
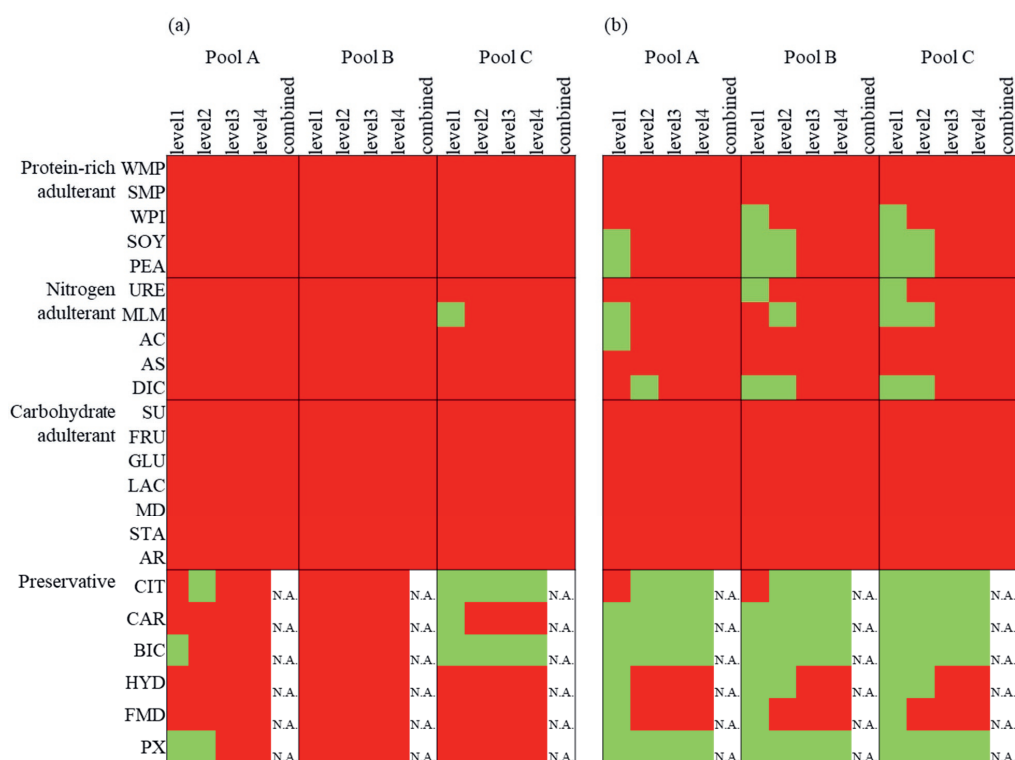
Fig. 4.2. The results of the adulterant test sets for the three milk pools based on (a) the measured boundaries; and (b) the variance-adjusted boundaries, indicating the potential to identify suspected milk adulterations. The samples with all results within the boundaries are coloured green, while the rest are coloured red. The full names of the adulterant are shown in the abbreviation list. N.A., not applicable.

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 4.2.5.2). 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. 4.2b. 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.

4.3.3.2. 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 performance in both scenarios (Table 4.4), hence, more detailed results are presented for SIMCA model (Fig. 4.3). 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 4.4), 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. 4.2). However, with the multivariate comparison, eight cases are missed (Fig. 4.3). Similarly for the variance-adjusted set, 24 are missed using the univariate approach (Fig. 4.2) and 40 using the multivariate approach (Fig. 4.3). Obviously, sensitivity may improve with larger numbers of samples, but currently the detection of adulteration seems to work better when using individual compositional parameters.

4.3.4. 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 4.5). 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 milk, and therefore the economic implication due to water dilution was not included in this section.

4

Table 4.4. 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 with $\gamma = 0.0001$ (based on measured dataset), and $\gamma = 0.1$ (based on variance-adjusted dataset).
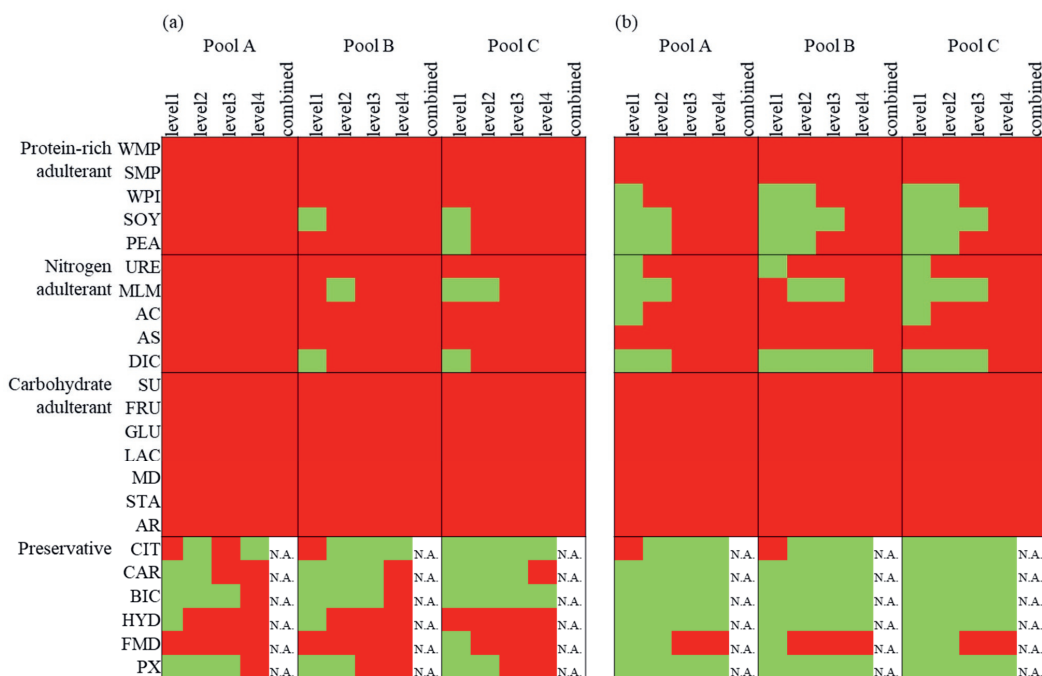
Fig. 4.3. The results of the adulterant test sets for the three milk pools based on the threshold of the of soft independent modelling of class analogies (SIMCA) model developed from (a) the measured dataset; and (b) the variance-adjusted dataset, indicating the classification of milk adulterations. 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. N.A., not applicable.

Table 4.5. 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	-	-

[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. 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.

As shown in Table 4.5, 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 4.3 and 4.5). 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.

4.3.5. Will the profitable adulterations be detected?

As discussed in section 4.3.4, 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 4.3.3, 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.

Acknowledgement

Declaration of competing interest

None.

4

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.

AOAC. (2019). Standard method performance requirements for whey protein: Casein ratio in infant formula. *Journal of AOAC International*, *96*(3), 501-501.

Beebe, K. R., Pell, R. J., & Seasholtz, M. B. (1998). *Chemometrics: A practical guide*. New York, NY: 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.

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.

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.

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.

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.

Finete, V. L. M., Gouvêa, M. M., Marques, F. F. 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.

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*. Berlin, Germany: Springer.

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*. Waltham, MA: 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.

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.

Hineno, M. (1977). Infrared spectra and normal vibration of β-d-glucopyranose. *Carbohydrate Research, 56*(2), 219-227.

International Organization for Standardization [ISO]. (2009). Milk — Determination of freezing point — Thermistor cryoscope method (ISO standard No. 5764). Retrieved 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). Retrieved 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). Retrieved 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.

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.

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.

Kačuráková, M., & Wilson, R. H. (2001). Developments in mid-infrared FT-IR spectroscopy of selected carbohydrates. *Carbohydrate Polymers, 44*(4), 291-303.

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., Clark, B., Home, R., & Frewer, L. (2019). Chinese consumer's attitudes, perceptions and behavioural responses towards food fraud. *Food Control*, *95*, 339-351.

4

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.

Liu, J., Ren, J., Liu, Z., & Guo, B. (2015). A new comprehensive index for discriminating adulteration in bovine raw milk. *Food Chemistry, 172*(Supplement C), 251-256.

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.

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.

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.

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.

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.

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.

Qlip. Payment criteria analyses. Retrieved from https://www.qlip.nl/en/component/content/article?id=358:payment-criteria-analyses.

Rattray, W., & Jelen, P. (1996). Protein standardization of milk and dairy products. *Trends in Food Science & Technology, 7*(7), 227-234.

Royal Friesland Campina. (2019). Milk price explained. Retrieved from 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.

Singh, P., & Gandhi, N. (2015). Milk preservatives and adulterants: Processing, regulatory and safety issues. *Food Review International, 31*(3), 236-261.

Sneddon, N., Lopez-Villalobos, N., Hickson, R., & Shalloo, L. (2013). Review of milk payment systems to identify the component value of lactose. *Proceedings of the New Zealand Society of Animal Production, 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.* Berlin, Germany: Springer.

Spink, J., Moyer, D. C., & Speier-Pero, C. (2016). Introducing the food fraud initial screening model (FFIS). *Food Control, 69*, 306-314.

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.

USDA-NASS. (2019). USDA national agricultural statistics service. Retrieved from https://www.nass.usda.gov/.

Yang, Y., Huisman, W., Hettinga, K. A., Liu, N., Heck, J., Schrijver, G. H., Gaiardoni, L., & van Ruth, S. M. (2019). Fraud vulnerability in the Dutch milk supply chain: Assessments of farmers, processors and retailers. *Food Control, 95*, 308-317.

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.

Zuivelverordening (2000). Uitbetaling van boerderijmelk naar kwaliteit, samenstelling en gewicht. Retrieved from https://wetten.overheid.nl/BWBR0010719/2005-01-01/0/HoofdstukI.

4

Supplementary material

Table S4.1. Information of adulterants purchased from the local shops.

Adulterants	Brand	Producer	Location of producer	Protein (% w/w)	Fat (% w/w)	Carbo-hydrates (% w/w)
Whole milk powder	TwoCows	Unidex BV	Hillegom, NL	25.7	28.2	36.5
Skimmed milk powder	Campina elk	Friesland Campina	Amersfoort, NL	35.4	0.6	50.5
Whey protein isolate	Royal Green	Royal Green	Hoorn, NL	93.0	0.3	2.5
Soy protein isolate	Pulsin	Pulsin	Gloucester, UK	90.0	3.0	0.5
Pea protein isolate	Pulsin	Pulsin	Gloucester, UK	81.7	1.7	0.8
Arrowroot powder	Smaakt	Smaakt	Ulvenhout, NL	0.3	0.1	85.0

Table S4.2. The result of the milk composition and feature of the control samples for the measured dataset [a].

Sample ID	Protein (% w/w)	Fat (% w/w)	Lactose (% w/w)	FPD	TS (% w/w)	SNF (% w/w)	Density (g/L)
1	3.67	3.82	4.63	0.569	13.73	9.88	1036
2	3.67	3.70	4.74	0.576	13.71	10.00	1036
3	3.78	3.72	4.74	0.581	13.86	10.13	1037
4	3.82	3.76	4.77	0.588	13.98	10.20	1037
5	3.49	3.76	4.60	0.552	13.44	9.64	1035
6	3.64	3.70	4.70	0.568	13.65	9.92	1036
7	3.28	3.75	4.57	0.548	13.17	9.38	1034
8	3.45	3.63	4.63	0.559	13.29	9.63	1035
9	3.55	4.03	4.62	0.558	13.80	9.74	1035
10	3.51	3.88	4.64	0.553	13.63	9.71	1036
11	3.47	3.87	4.68	0.566	13.62	9.72	1036
12	3.47	3.64	4.68	0.553	13.38	9.71	1036
Pool A	3.74	3.77	4.75	0.576	13.88	10.09	1037
Pool B	3.52	3.68	4.61	0.561	13.44	9.68	1036
Pool C	3.50	3.84	4.65	0.562	13.58	9.71	1035

[a] ID: Identity; FPD: freezing point depression; TS: total solids; SNF: solids non-fat.

Table S4.3. The result of the milk composition and feature of the control samples for the variance-adjusted dataset [a].

Sample ID	Protein (% w/w)	Fat (% w/w)	Lactose (% w/w)	FPD	TS (% w/w)	SNF (% w/w)	Density (g/L)
1	3.76	3.87	4.59	0.574	13.84	9.94	1036
2	3.77	3.62	4.81	0.588	13.81	10.18	1036
3	3.98	3.66	4.82	0.597	14.10	10.44	1038
4	4.06	3.75	4.88	0.611	14.35	10.59	1038
5	3.40	3.74	4.53	0.539	13.27	9.47	1034
6	3.70	3.63	4.73	0.572	13.68	10.02	1036
7	2.98	3.72	4.48	0.531	12.73	8.94	1031
8	3.32	3.49	4.59	0.554	12.97	9.44	1033
9	3.53	4.28	4.58	0.552	13.99	9.67	1034
10	3.45	3.99	4.62	0.541	13.65	9.61	1035
11	3.37	3.97	4.70	0.568	13.62	9.62	1035
12	3.37	3.50	4.69	0.542	13.15	9.60	1036
Pool A	3.91	3.76	4.84	0.588	14.14	10.37	1038
Pool B	3.47	3.59	4.56	0.558	13.27	9.55	1035
Pool C	3.42	3.91	4.63	0.559	13.55	9.60	1033

[a] FPD: freezing point depression; TS: total solids; SNF: solids non-fat.

4

Table S4.4. The liner model performance comparing apparent results and the spiked levels, for the parameters of total solid, solid non-fat and density for each adulterant [a].

Category	Adulterant	Total solids		Solids non-fat		Density	
		m	*R²*	*m*	*R²*	*m*	*R²*
Protein-rich adulterants	WMP	0.97	1.00	0.76	1.00	2.92	0.99
	SMP	0.68	0.99	0.80	0.99	3.25	0.99
	WPI	0.31	0.99	0.35	0.97	0.62	0.69
	SOY	0.22	0.93	0.25	0.97	0.67	0.94
	PEA	0.25	0.99	0.29	0.98	0.98	0.96
Nitrogen-based adulterants	URE	-0.24	0.99	-0.13	0.96	0.90	0.95
	MLM	0.03	0.19	0.07	0.76	1.47	0.96
	AC	-0.44	0.99	-0.33	1.00	0.22	0.47
	AS	-0.26	0.99	-0.10	0.96	8.63	1.00
	DIC	0.16	0.93	0.20	0.95	0.40	0.49
Carbohydrate-based adulterants	SU	0.95	0.99	1.16	0.99	-1.15	0.88
	GLU	1.01	1.00	1.14	0.99	4.55	1.00
	FRU	0.92	1.00	1.13	0.99	5.42	0.99
	LAC	0.69	1.00	0.84	0.99	5.23	1.00
	MD	0.94	0.99	1.22	0.99	2.58	0.99
	STA	0.90	0.98	0.44	0.97	-3.42	0.94
	ARR	1.07	1.00	0.53	0.99	-3.73	0.99
Preservatives	CIT	0.00	0.26	0.02	0.39	0.18	0.40
	CAR	0.01	0.55	0.05	0.75	0.12	0.21
	BIC	0.00	0.13	0.02	0.44	0.02	0.27
	FMD	0.04	0.54	0.09	0.81	0.17	0.37
	PX	0.03	0.54	0.05	0.50	0.05	0.62
	HYD	-0.04	0.37	0.05	0.36	0.05	0.29

[a] The slope (*m*) and R-square (*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; 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.

4

CHAPTER 5

Exploration of an ultrasonic pulse echo system for comparison of milks, creams, and their dilutions

Abstract

Rapid measurements are valuable in food authentication. As a first step towards application of sound for fast food authentication, the influence of fat content and dilution of liquid milk/cream products on their ultrasound characteristics was examined. A total of 18 liquid dairy products, ranging in fat content from 0 to 60 g/100 g product, including nine cream samples and nine milk samples, were diluted to four levels, which resulted in 90 samples. Ultrasonic velocity, together with physical properties including viscosity, density, and macronutrient composition of the tested dairy products, was determined. The results show that the ultrasonic velocity in the dairy products increased with a decreasing fat content of the samples. There was a significant positive correlation between the velocity and density, which was further related to the impact of the fat content. Water dilution caused different changes in velocity, depending on the ratio of fat content and continuous phase of the samples. A higher fat content of the samples resulted in a smaller effect of dilution on velocity. This study showed that the ultrasonic system allows rapid and non-destructive analysis for comparison of liquid dairy samples and appears promising for detection of dilution.

Keywords

Composition; Dairy; Density; Rheology; Ultrasonic velocity

5.1. Introduction

Animals have utilized low intensity ultrasound for millions of years, for instance bats and dolphins use echo location for hunting or navigation. Ultrasound is defined as sound waves with frequencies exceeding 20 kHz, which is above the hearing limit of the human ear (Awad, Moharram, Shaltout, Asker, & Youssef, 2012). Sound propagates through materials as mechanical waves, and its velocity and intensity are impacted by the interaction between the sound wave and materials. Consequently, analysing the velocity and the change of intensity could help to obtain information about the some of the properties of materials. For instance to understand bone density (Lee, Coan, & Bouxsein, 1997), to evaluate compressive strength of concrete (Demirboğa, Türkmen, & Karakoç, 2004), etc.

In modern times, ultrasonic techniques have also been applied by food industries. These industrial applications for food characterization include, amongst others, fat content analysis of edible oil (Yan et al., 2019), determination of the botanical origin of sugars in juices (Contreras, Fairley, McClements, & Povey, 1992), and assessment of the composition of fish and poultry (Sigfusson, Decker, & McClements, 2001). As a rapid screening approach without any preparation or destruction of the sample, ultrasonic techniques could be easily fitted in online monitoring systems and allow for real-time adjustments during processing. Among several parameters which could be determined by ultrasound, measurement of the ultrasonic velocity is considered as the basis of the ultrasonic techniques used to evaluate food properties (McClements & Gunasekaran, 1997).

One of the food categories for which ultrasonic techniques have been used before are dairy products. Dairy products are consumed worldwide. There are various types of dairy products being produced, including liquid as well as solid products. Liquid milk and cream are two principal liquid dairy products. The cost of these dairy products is related to their composition properties such as their fat, protein, and lactose contents. It is therefore important for manufacturers to be able to characterize these products and adjust them. One area of characterization is to detect adulteration. As an important food in the human diet with a high nutritional value, milk is a constant target of adulteration (Nascimento, Santos, Pereira-Filho, & Rocha, 2017). Despite the various possible types of milk adulteration, dilution is the easiest way to adulterate milk, especially if milk payment is primarily based on the volume of milk. The commonly used techniques for the measurement of milk composition range from cheap and high-throughput techniques like Fourier transformed infrared spectroscopy (FTIR), to more expensive and labour-intensive ones like liquid chromatography and mass spectrometry. Nevertheless, these techniques are either time consuming or destructive. Some scholars reported applications of the ultrasonic technique to dairy products, for instance, for evaluating the fat content and particle size distribution, using ultrasonic attenuation (Dukhin, Goetz, & Travers, 2005), monitoring the crystallization of milk fat (Singh, McClements, & Marangoni, 2004), testing reconstitution of milk powders (Meyer, Rajendram, & Povey, 2006), monitoring the phase change for yogurt fermentation (Ogasawara, Mizutani, Ohbuchi, & Nakamura, 2006), and detection of milk adulterated with melamine (Elvira, Rodríguez, & Lynnworth, 2009).

5

These previous studies indicate that ultrasonic velocity is influenced by the particles and fat content in liquid samples., However, no specific study was found on profiling of various types of dairy emulsions, i.e. milks and creams of different fat contents, using ultrasonic velocity, it is and therefore is promising to apply this for rapid and non-invasive technique in characterization of dairy products.

In this study, we explored the influence of fat content of liquid milk/cream products on their ultrasound characteristics. A pulse-echo ultrasonic system was applied to characterize diverse liquid dairy products with a wide range of fat contents, from double creams to skimmed milks. The range was further extended by dilution of the products to four different levels, to test for the sensitivity towards detecting this common type of adulteration. The study thus aims to evaluate the developed ultrasonic velocimetry system for profiling liquid dairy products, and to examine the underlying causes for the variation in ultrasonic velocity between the various products.

5.2. Materials and methods

5.2.1. Samples

A total of 18 liquid dairy product samples of four different brands were purchased from supermarkets in Ireland during summer (June to August 2018). They included six types of dairy products of which three samples were obtained each: double cream samples, normal cream samples, and light cream samples, as well as whole milk samples, semi skimmed milk samples, and skimmed milk samples. All samples were stored at 4 °C and analysed before their expiry date.

These 18 samples were then mixed with demineralised water in the following water/dairy sample ratios (w/w) to create four dilution levels for each sample, resulting in a water/product ratio 20/80, 40/60, 60/40, and 80/20, based on weight percentage. The mixed solution was stirred for 1 minute and was measured right after stirring. The final sample set included 90 samples, including 18 undiluted samples and 72 diluted samples.

5.2.2. Ultrasonic velocity analysis

A purpose-built pulse-echo system was utilized to measure the ultrasonic velocity. As shown in Fig. 5.1, the system comprised a sample cell with stainless-steel bottom, an immersion transducer (Panametrics-NDT, Olympus NDT U.K. Ltd, Rotherham, South Yorkshire, UK) with a central frequency of 5 MHz, a pulse generator/receiver (Panametrics-NDT Model 5800, Olympus NDT U.K. Ltd., Rotherham, South Yorkshire, UK), an oscilloscope (Tektronix TDS 210, Tektronix UK Ltd., Bracknell, UK), and a computer (Dell, Texas, US).

The principle of the pulse echo system has been described before (Awad et al., 2012; Yan et al., 2019). The monopolar electrical pulses are generated/received by a computer-controlled pulse generator/receiver, and then transformed to ultrasonic pulses by the transducer. The pulse

generator/receiver was operated in pulse echo mode, with a pulse repetition frequency of 1 kHz, 12.5 µJ pulse energy with 20 dB input attenuation, and 20 dB receiver amplification. The ultrasonic pulse was generated from the transducer, propagated through the dairy samples, bounced back from the stainless-steel bottom of the sample cell (solid reflector), and propagated back to the pulse receiver. The transducer then transformed the returned ultrasonic pulses to electrical signals, which were shown on the oscilloscope. Next, the electrical signals were transferred to the computer from the oscilloscope. The measurements were conducted with various distances between the transducer and the bottom of the sample cell. The transducer was mounted on a micrometre stage with a positioning accuracy of $\pm$ 0.005 mm. The signals were collected using a MATLAB (R2015b, MathWorks Inc., Natick, MA, USA) program.
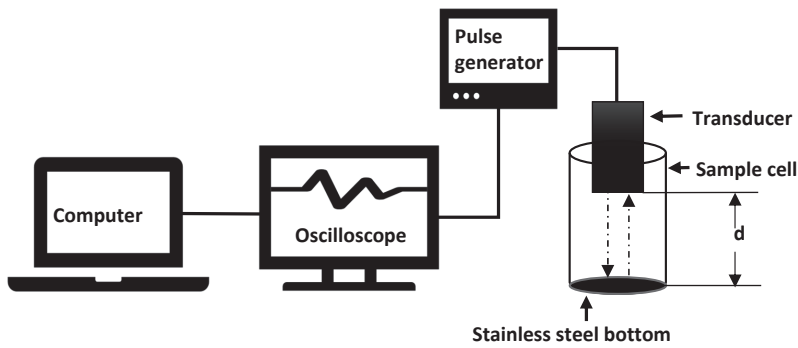

Fig. 5.1. Schematic diagram of the ultrasonic pulse echo measurement device.

The ultrasonic velocity was calculated using the distances and the propagation times of the echoes. Six consecutive distances (10.00 $\pm$ 0.005, 12.00 $\pm$ 0.005, 14.00 $\pm$ 0.005, 16.00 $\pm$ 0.005, 18.00 $\pm$ 0.005, and 20.00 $\pm$ 0.005 mm) were used for each sample measurement. The ultrasonic velocity was then determined by two measurements of two consecutive distances using Eq. (5.1):

$$c = \frac{2 \times (d_2 - d_1)}{t_2 - t_1} \hspace{3cm} \text{Eq. (5.1)}$$

Where c is the ultrasonic propagating velocity (m/s) in the sample; d_1 is the shorter distance of the one-way echo path (m), d_2 is the consecutive distance of the one-way echo path (m); t_1 is the time of receiving the first signal ($\pm$ 0.1 µs) under the distance of d_1, t_2 is the time of receiving the first signal ($\pm$ 0.1 µs) under the distance of d_2. Five velocity values were determined based on the six measurements and averaged. This resulted in a dataset of 18 samples x five dilution levels x six measurements is 540 data points. All samples were placed in water bath to maintain the temperature (24.0 $\pm$ 0.1°C) before the measurements. The measurements were performed at 24.0 $\pm$ 0.5°C, which were monitored by an electronic thermometer. The device was calibrated by measuring the ultrasonic velocity in water (1494 $\pm$ 3 m/s at 24.0 °C) (Engineering ToolBox, 2004).

5.2.3. Viscosity analysis

A stress-controlled rheometer (MCR 301, Anton Paar, Austria) with a stainless-steel double gap geometry (DG26.7/TI-SN21833) was used to measure the viscosity. The viscosity was measured by increasing the shear rate from 0.1 to 200 /s in 3 min. A total of 200 data points was recorded for each sample. All experiments were performed in duplicate at 24.0 ± 0.5°C. The viscosity measured at the shear rate 100 /s was used for the statistical analysis.

5.2.4. Density analysis

The density of the 18 commercial samples was determined using a gravimetric method. A measuring cylinder (25.0 mL) was weighed (m_1) using an electronic balance (± 0.001 g), then 25.0 ± 0.1 mL of sample (v) was quantified using the measuring cylinder. Subsequently, the measuring cylinder and the sample were weighed together (m_2). The density of the sample was calculated by Eq. (5.2):

$$\rho = \frac{m_2 - m_1}{v} \times 1000 \qquad\qquad \text{Eq. (5.2)}$$

Where ρ is the density of the sample (g/L), m_1 is the weight of empty measuring cylinder (g), m_2 is the total weight of the measuring cylinder and the cream sample (g), and v is the accurate volume of the sample (mL). The sample was equilibrated at 24.0 ± 0.1 °C in a water bath before the measurement, and the measurements were conducted in duplicate at room temperature (24.0 ± 0.5 °C).

5.2.5. Compositional analysis

The composition of the 18 commercial samples was measured using a MilkoScan FT120 instrument (Foss Electric, Hilleroed, Denmark). The compositional features obtained for the milk samples included fat content, protein content, total solids content (TS), solids non-fat content (SNF), and lactose content, and for the cream samples included the fat content, protein content, TS content, and SNF content. Lactose could not be measured in the creams because the cream module of the instrument was not calibrated for this parameter. Samples were equilibrated to 24.0 ± 0.1 °C in a water bath prior to analysis. All analyses were conducted in duplicate, resulting in 36 measurements. The average value of the duplicate was used for analysis.

5.2.6. Calculations and statistical analysis

5.2.6.1. Density and compositional calculations

The density and composition of the diluted samples were calculated based on their non-diluted counterparts. The density and compositional feature of the *a%* sample was calculated using Eq. (5.3) and Eq. (5.4), respectively:

$$\rho_{a\%} = \frac{100\,g}{[a\% \times 100\,g\,/\,\rho_{dairy}] + [(1-a\%) \times 100\,g\,/\,\rho_{water}]} \qquad \text{Eq. (5.3)}$$

$$X_{a\%} = a\% \times X_{dairy} + (1 - a\%) \times X_{water} \qquad \text{Eq. (5.4)}$$

Where $\rho_{a\%}$ is the density of the diluted sample (g/L), $a\%$ is equal to the weight percent of product in the diluted sample, thus being 80%, 60%, 40%, and 20%, respectively. ρ_{dairy} is the density of the sample (g/L) , ρ_{water} is the density of water (g/L). $X_{a\%}$ is the compositional feature of the diluted sample (g/100 g), X_{dairy} is the compositional feature of the commercial dairy product (g/100 g), i.e. protein, fat, lactose, TS, SNF for the milk samples, and protein, fat, TS, SNF for the cream samples. X_{water} is the corresponding feature of water (g/100 g).

5.2.6.2. Calculation of compressibility

The compressibility of the samples is calculated based on the means of velocity and density. The velocity in a fluid is described by the Wood equation (Wood, 1946), which is also called Laplace equation:

$$v = \sqrt{\frac{1}{k \times \rho}} \qquad \text{Eq. (5.5)}$$

Consequently the compressibility was obtained using Eq. (5.6):

$$k = \frac{1}{v^2 \times \rho} \qquad \text{Eq. (5.6)}$$

Where k is the compressibility (Pa^{-1}), v is the velocity of ultrasound (m/s), and ρ is the density (g/L).

5.2.6.3. Statistical analysis

Means and standard deviations of the measured physical properties were calculated for each sample, and for the six dairy product groups. Because not all the data was normally distributed, non-parametric Kruskal-Wallis tests were applied for comparisons between groups, followed by the pairwise Mann-Whitney U tests. Correlation between physical properties were determined by computing Pearson correlation coefficients (*r*). The non-parametric tests were performed using SPSS statistic 25 (IBM, Chicago, IL, USA), whereas the other analyses were performed using R 3.6.1 (R Foundation for Statistical Computing, Vienna, Austria).

5

5.3. Results and discussion

5.3.1. Ultrasonic velocity in the dairy products

5.3.1.1. Ultrasonic velocity in the original dairy samples

The ultrasonic velocities in the 18 dairy samples are presented in Table 5.1. In general, the ultrasonic velocity in the milk samples (ranging from 1520 to 1525 m/s) is significantly higher than that in the cream samples (ranging from 1489 to 1506 m/s). More specifically, the ultrasonic velocity in the light cream samples was significantly higher than those in the double cream and normal cream samples, and the ultrasonic velocity in the milk samples appears to have an upward trend for the milk products from whole milk samples to skimmed milk samples (Table 5.1). Based on this data, it may be hypothesised that the ultrasonic velocity in dairy products depends on its fat content.

Table 5.1. Means and standard deviations of the ultrasonic velocity, viscosity [a] and density for the six dairy groups and 18 commercial dairy samples. Different letters indicate significant differences between the dairy groups (Kruskal-Wallis test followed by pairwise Mann-Whitney U test, $p < 0.05$).

Sample	Velocity (m/s)	Viscosity (mPa.s)	Density (g/L)
Double cream 1	1491.4 ± 6.1	30.0 ± 0.5	974 ± 2
Double cream 2	1493.7 ± 7.3	17.0 ± 0.7	993 ± 2
Double cream 3	1483.7 ± 3.0	36.6 ± 1.7	981 ± 1
Mean double cream group	1489.6 ± 5.2 a	27.9 ± 10.0 ab	983 ± 9 a
Normal cream 1	1499.3 ± 4.7	11.5 ± 0.3	995 ± 2
Normal cream 2	1497.0 ± 2.5	11.8 ± 0.2	998 ± 2
Normal cream 3	1492.5 ± 0.0	15.1 ± 0.7	980 ± 3
Mean normal cream group	1496.3 ± 3.4 a	12.8 ± 2.0 c	993 ± 12 a
Light cream 1	1507.2 ± 3.1	86.4 ± 10.5	1013 ± 1
Light cream 2	1503.8 ± 0.0	6.1 ± 0.0	1007 ± 1
Light cream 3	1508.3 ± 2.5	28.6 ± 0.3	1013 ± 0
Mean light cream group	1506.4 ± 2.4 b	40.4 ± 41.4 a	1011 ± 3 b
Whole milk 1	1520.9 ± 4.1	2.2 ± 0.0	1036 ± 0
Whole milk 2	1519.8 ± 4.8	2.0 ± 0.0	1035 ± 1
Whole milk 3	1523.2 ± 3.2	2.2 ± 0.0	1037 ± 0
Mean whole milk group	1521.3 ± 1.8 c	2.1 ± 0.1 bc	1036 ± 1 c
Semi skimmed milk 1	1522.1 ± 2.6	1.8 ± 0.0	1037 ± 1
Semi skimmed milk 2	1522.1 ± 4.8	1.9 ± 0.0	1037 ± 0
Semi skimmed milk 3	1525.6 ± 2.6	1.9 ± 0.0	1038 ± 1
Mean semi skimmed milk group	1523.2 ± 2.0 c	1.8 ± 0.1 bc	1037 ± 1 cd
Skimmed milk 1	1525.6 ± 4.9	2.3 ± 0.1	1040 ± 1
Skimmed milk 2	1523.2 ± 3.2	1.7 ± 0.0	1042 ± 1
Skimmed milk 3	1526.7 ± 4.1	1.7 ± 0.0	1038 ± 1
Mean skimmed milk group	1525.2 ± 1.8 c	1.9 ± 0.4 c	1040 ± 2 d

[a] Viscosity at shear rate 100 /s was used.

Milk and cream can both be considered as an emulsion of continuous water phase with a dispersed lipid phase (Huppertz & Kelly, 2006). The main content of the continuous phase, also called milk plasma, is water containing protein, lactose, minerals, and other dissolved solids. The dispersed phase is composed of milk fat globules. Ultrasonic velocity in a pure liquid was first described by the Wood equation (Wood, 1946). Based on the equation, scholars developed the theory that the ultrasonic velocity in an emulsion depends on the volume ratio of the two phases of emulsion, the dissolved solid content in the continuous phase, and the physical properties, for instance the droplet size, of the fat fraction of the emulsion (McClements & Povey, 1987; Povey, 1989). The dissolved solids in the continuous phase of an emulsion correlated positively to the ultrasonic velocity, while the volume fraction of milk fat was shown to have a negative effect on the ultrasonic velocity. Considering the composition of the milk plasma is similar for milk and cream products, therefore the large difference in the ratio of the continuous phase and dispersed phase is probably a key factor affecting the ultrasonic velocity. This is consistent with the results that ultrasonic velocity is decreasing with the increasing fat content of samples, as shown by the result of the ultrasonic velocities in this study (Table 5.1).

To sum up, the ultrasonic velocity increased with decreasing fat content of the samples. Using the ultrasonic velocity, cream samples can be differentiated from milk samples, and light creams can be differentiated from double and normal creams. Considering that the composition of continuous milk plasma phase for the six dairy products is almost identical, the proportion of fat content appears to be the main factor determining the difference in ultrasonic velocity between the six product groups.

5.3.1.2. Ultrasonic velocity in the diluted dairy samples

The ultrasonic velocity in the diluted liquid dairy samples revealed a variable response that differed between the six dairy products (Fig. 5.2). For the three liquid milk samples (whole milk, semi skimmed milk, and skimmed milk), the ultrasonic velocity decreased and got closer to the velocity of water (1492.5 m/s) as the sample was more diluted. The velocity in the diluted light cream samples showed a similar trend with those in the diluted milk samples, but with a less steep slope. The change of velocity in the diluted normal and double cream samples was much smaller.

It is important to point out that the dilution of liquid dairy samples with water could cause two opposite effects on the ultrasonic velocity: on the one hand, lowering the fat fraction by dilution could cause an increase in velocity (McClements, 1988). On the other hand, lowering the dissolved solids content in the continuous phase could cause a decrease of velocity (McClements & Gunasekaran, 1997). Since the milk plasma is the principal phase of the milk products (> 95% weight fraction), the declining dissolved solids concentration of these samples after dilution is considered as the principal factor for the velocity decrease. The velocity in the light cream generally declined less with dilution, compared with the milk samples (Fig. 5.2), which is probably caused by the smaller difference between the milk plasma and fat content in the light cream than in the milk samples. Consequently, the velocity decline caused by dilution

of the dissolved milk plasma was partly counteracted by the simultaneously decreasing fat content. This effect became more extensive in the normal and double cream samples, which means that the impact of dilution of the continuous milk plasma phase was almost completely balanced out by the decreasing fat content, resulting in almost no change of the velocity in these diluted cream samples (Fig. 5.2).
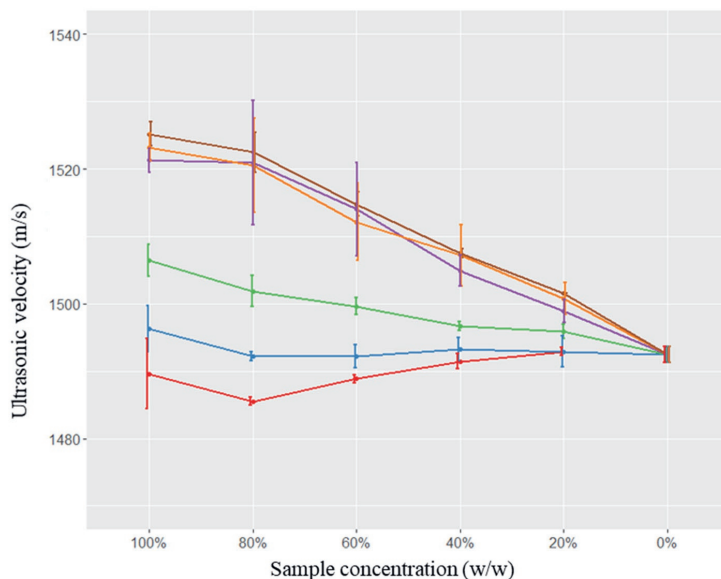

Fig. 5.2. Ultrasonic velocities of in the dairy samples double cream (red), normal cream (blue), light cream (green), whole milk (purple), semi skimmed milk (orange), and skimmed milk (brown) of various concentrations (100%-0%).

In summary, the ultrasonic velocity in the diluted milk samples decreased with dilution level, caused by a reduced total dissolved content in the samples. The ultrasonic velocity in the diluted cream samples showed smaller impact by the increased level of dilution, due to the counterbalancing effect of dilution on the dissolved milk plasma phase and the fat content.

5.3.2. Viscosity of the dairy products

The shear rheological behaviour of the undiluted samples is shown in Fig. 5.3. The apparent viscosities of the cream samples were generally larger than those of the milk samples. For emulsion systems, the viscosity increases linearly with the fraction of the dispersed phase (Bakshi & Smith, 1984; Huppertz & Kelly, 2006; Kyazze & Starov, 2004; Phillips, Mcgiff, Barbano, & Lawless, 1995), explaining why the larger fat fraction of the cream samples correlated with the higher viscosity (Fig. 5.3). This is also in line with the result that higher viscosity was observed for the whole milk samples compared to the semi skimmed and skimmed milk.

In contrast to the general relation between fat content and viscosity, it is noted that the viscosity of the light cream samples was higher, and its decreasing rate with the shear rate was lower, compared with the other two types of cream products (Fig. 5.3). This may be due to the fact that, for all the 18 commercial products used in this study, only the three light cream samples contained ingredients other than milk, including carrageenan, skimmed milk powder and glucose (according to their labels). These components are added as thickeners to obtain the desired texture of the product, which is more similar to higher fat cream products. Previous studies showed that carrageenan can greatly increase the viscosity of a cream product, by means of bridging the casein micelles (Camacho, Martínez-Navarrete, & Chiralt, 2005; Langendorff et al., 2000). As a result, the viscosity of the light cream samples was higher than was expected based on its fat content.

To summarize, the viscosity of the cream samples was higher than the milk samples. Due to the addition of polymers, the viscosity of light creams was higher than the value expected from its fat content only.

Fig. 5.3. The result of the viscosity in the double cream (red), normal cream (blue), light cream (green), whole milk (purple), semi skimmed milk (orange), and skimmed milk (brown) samples. The average value of the three samples in each group was used.

5.3.3. Density of the dairy products

The density of the six groups of dairy products showed an upward trend with decreasing fat contents (Table 5.1 and 5.2). The density of the cream products was significantly lower than that of the milk products. As discussed in 5.3.1.1, the dairy products could be considered as an emulsion of two phases: milk plasma and milk fat globules. The density of the milk plasma changes very slightly depending on the components dissolved in it, which results in a density

of around 1.030 g/L at 20 °C. The density of the milk fat depends on variations in the fat composition, resulting in a density around 0.92 g/L at 20 °C (Walstra, Wouters, & Geurts, 2005). The lower density of the milk fat therefore causes a decreasing density of dairy products with increasing fat content, leading to the highest density for the skim milk and the lowest density for the double cream.

Table 5.2. Means and standard deviations of the milk composition for the six dairy groups ($n = 3$).

Features	Double cream	Normal cream	Light cream	Whole milk	Semi skimmed milk	Skimmed milk
Fat (g/100 g)	58.0 ± 3.1	50.3 ± 0.9	35.7 ± 2.4	3.5 ± 0.2	1.6 ± 0.6	0.2 ± 0.1
Protein (g/100 g)	3.3 ± 0.1	3.8 ± 0.1	4.5 ± 0.5	3.9 ± 0.1	3.9 ± 0.1	4.0 ± 0.12
TS (g/100 g)	63.3 ± 2.5	56.3 ± 0.7	42.8 ± 2.3	13.6 ± 0.3	11.8 ± 0.5	10.6 ± 0.6
SNF (g/100 g)	0.6 ± 0.3	1.7 ± 0.2	2.7 ± 0.5	10.0 ± 0.1	10.1 ± 0.2	10.4 ± 0.5
Lactose (g/100 g)	N.A. *	N.A.	N.A.	4.5 ± 0.0	4.5 ± 0.1	4.7 ± 0.3

* N.A., non-available because the feature was not calibrated for the instrument for the specific products.

5.3.4. Relations between the ultrasonic velocity, viscosity, and density

5.3.4.1. Relation between the ultrasonic velocity, viscosity, and density for the original samples

To explore the relation of the ultrasonic velocity, viscosity, and density for the original and diluted dairy samples, scatter plots between the velocity and viscosity and density for both original and diluted dairy samples are presented (Fig. 5.4). As shown in Fig. 5.4a, no clear relationship between the velocity and viscosity was observed for the original dairy samples, which was supported by the Pearson correlation analysis, which resulted in no significant correlations between these two features for the original dairy samples ($r = -0.45$, $p = 0.063$). On the contrary, it can clearly be observed that the velocity in the dairy samples increased with their density (Fig. 5.4b). Correspondingly, a significant positive correlation between velocity and density ($r = 0.98$, $p < 0.05$) for the original dairy samples was obtained.

5.3.4.2. Relation between the ultrasonic velocity, viscosity, and density for the diluted dairy samples

It is shown in Fig. 5.4a that no clear pattern was observed between the velocity and viscosity for the diluted dairy samples. The velocity in the diluted samples increased with their density, as shown in Fig. 5.4b, resulting in a significant positive correlation between the velocity and density ($r = 0.94$, $p < 0.05$). Considering that density depends largely on the composition, the correlation between velocity and density may be related to the composition of the samples. The relations between the velocity, density, and fat content were therefore analysed.

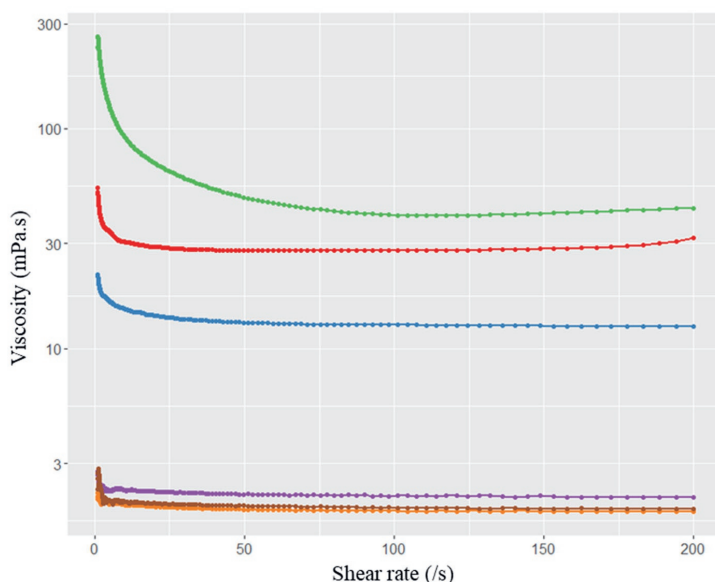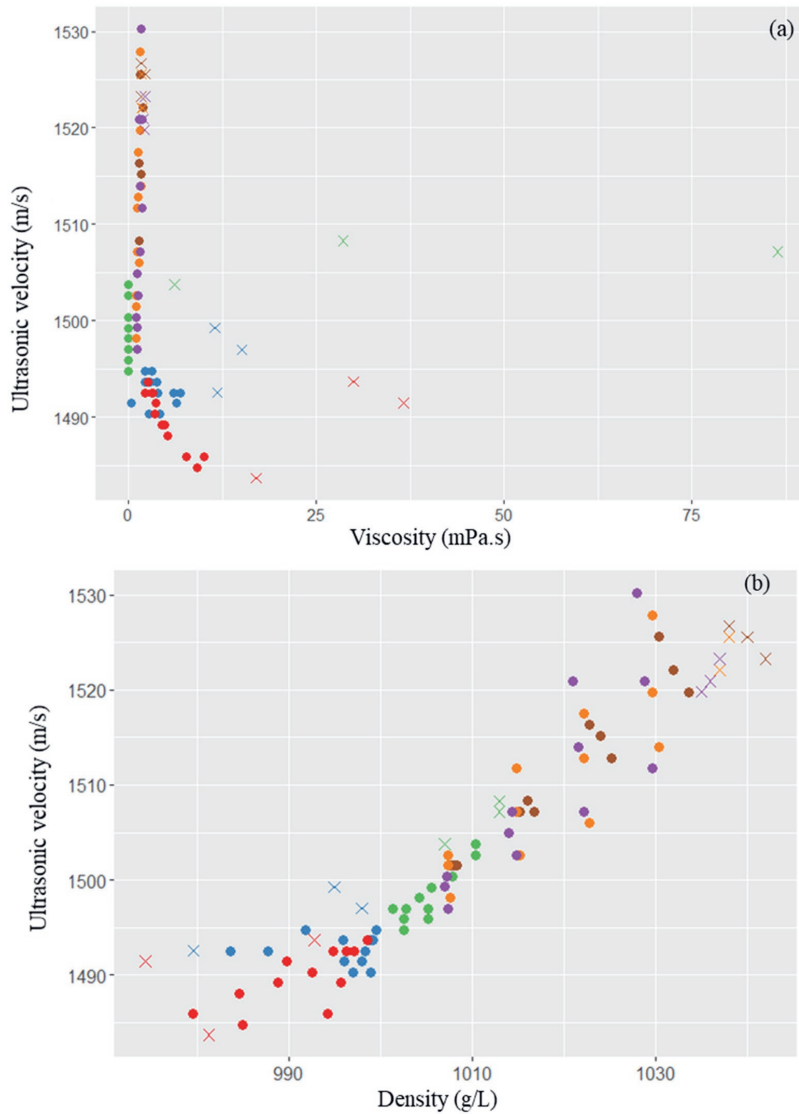**Fig. 5.4.** Scatter plots of (a) ultrasonic velocity and viscosity for the double cream (red), normal cream (blue), light cream (green), whole milk (purple), semi skimmed milk (orange), and skimmed milk (brown), for the original samples (x), and their diluted counterparts (circle); and (b) ultrasonic velocity and density for the double cream (red), normal cream (blue), light cream (green), whole milk (purple), semi skimmed milk (orange), and skimmed milk (brown), for the original samples (x) and their diluted counterparts (circle).

5.3.5. Relation between the ultrasonic velocity, density, compressibility, and fat content

To explore the correlation of the velocity, density, and the compositional features for the original dairy samples, the Pearson correlation coefficients (r) between the physical and compositional properties are presented in Table 5.3. The velocity and density were significantly

negatively correlated with the fat content and TS content, and positively correlated with the SNF content. For the original dairy samples, the fat to milk plasma ratio of the samples decreases, starting from the samples of highest fat content (double cream milk) to the ones containing no fat (skimmed milk). Therefore, the fat content is considered the main factor explaining the differences of the velocity in the samples. This is in line with the hypothesis that the ultrasonic velocity in the dairy products depends on the fat content of the samples.

Table 5.3. Pearson correlation coefficients for the correlation between the ultrasonic velocity, density, and the compositional features [a] for the original dairy group ($n = 18$).

Compositional features	Velocity	Density
Fat	-0.98 *	-0.98 *
Protein	0.43	0.45
Total solids	-0.98 *	-0.98 *
Solids non-fat	0.96 *	0.95 *

[a] Only the available compositional features for both milk and cream samples are used for the calculations.
* Significant correlations ($p < 0.05$) are marked.

The changes of the velocity caused by dilution of the six types of dairy products showed different trends, which were further related to the changes of the compressibility. Relations between the velocity, compressibility, and fat content in this study are presented in Fig. 5.5. According to the Wood equation (Eq. 5.5), velocity in a fluid is influenced by both density and compressibility. The dilution of the milk samples impacts differently on the density and compressibility, which resulted in opposite influences on velocity: on the one hand, the increasing compressibility led to a decrease in the velocity; on the other hand, the decreasing density led to an increase in the velocity. The results show that the velocity in the milk samples decreased with the dilution (Fig. 5.5b), indicating that the increase in compressibility has dominant effects on velocity, which resulted from the dilution of milk plasma. While for the diluted double cream and normal cream samples, the increase of velocity indicates that the decrease in compressibility has a larger impact, which resulted from the considerable reduction in fat content (Fig. 5.5). This result is in line with the discussion in section 5.3.1.2, that the different ratio of fat to milk plasma can affect the velocity in diluted samples: the larger the difference between the proportion of the two phases, the greater the change in velocity caused by dilution. The counteracting effects between the fat content and continuous milk plasma phase on the velocity provide a potential for determination of the ratios of the two phases in dairy products using ultrasonic velocity, as a larger effect of dilution indicates a larger fraction of milk plasma. In addition, for the milk samples with a relatively low-fat content, the velocity might be used for detection of adulteration with water.

In summary, the difference of fat content is the principal factor resulting in the difference of ultrasonic velocity in the undiluted dairy samples. While for the diluted samples, the changes in ultrasonic velocity caused by dilution are influenced by the ratio of fat and continuous phases in the original samples.
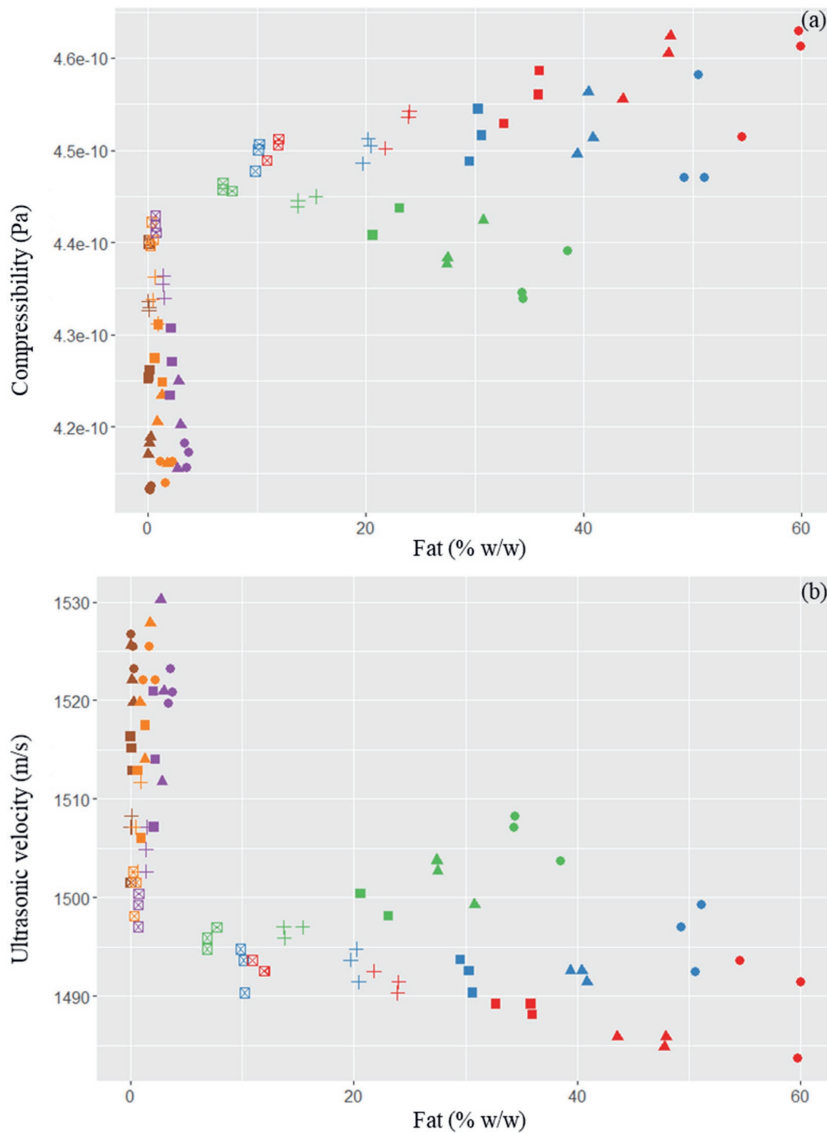
Fig. 5.5. Relation between (a) compressibility and fat content, and (b) ultrasonic velocity and fat content, for the double cream (red), normal cream (blue), light cream (green), whole milk (purple), semi skimmed milk (orange), and skimmed milk (brown) samples of concentration of 100 g/100 g (circle), 80 g/100 g (triangle), 60 g/100 g (square), 40 g/100 g (+), and 20 g/100 g (x in square).

5.4. Conclusions and outlook

The study showed a potential to apply ultrasonic velocity measurement for characterization of liquid dairy products. The ultrasonic velocity in liquid dairy products decreases with increasing fat content. The fat content appears the determining factor for ultrasonic velocity in undiluted dairy products, whereas both fat content and the fraction of the

continuous phase play a role in the changes in ultrasonic velocity in diluted dairy products. Consequently, the higher the fat content of a sample, the smaller the change in ultrasonic velocity caused by dilution.

The strong correlation between the velocity and fat content could be used for characterisation of the dairy products. The opposite effects on velocity caused by dilution, namely increase of velocity with decreasing fat content and decrease of velocity with dilution of the continuous phase, provides a potential for determination of the ratio of two phases in the dairy products using ultrasonic velocity. As a rapid and non-destructive technique, ultrasonic velocity is promising to be applied for detection of milk adulteration with water. In addition, it would be also interesting to explore its capability with regard to the detection of other common types of milk adulterations, especially those that are aimed at masking the dilution with water.

Acknowledgement

Declaration of competing interest

None.

References

Awad, T. S., Moharram, H. A., Shaltout, O. E., Asker, D., & Youssef, M. M. (2012). Applications of ultrasound in analysis, processing and quality control of food: A review. *Food Research International, 48*(2), 410-427.

Bakshi, A., & Smith, D. E. (1984). Effect of fat content and temperature on viscosity in relation to pumping requirements of fluid milk products. *Journal of Dairy Science, 67*(6), 1157-1160.

Camacho, M. M., Martínez-Navarrete, N., & Chiralt, A. (2005). Rheological characterization of experimental dairy creams formulated with locust bean gum (LBG) and λ-carrageenan combinations. *International Dairy Journal, 15*(3), 243-248.

Contreras, N. I., Fairley, P., McClements, D. J., & Povey, M. J. W. (1992). Analysis of the sugar content of fruit juices and drinks using ultrasonic velocity measurements. *International Journal of Food Science & Technology, 27*(5), 515-529.

Demirboğa, R., Türkmen, İ., & Karakoç, M. B. (2004). Relationship between ultrasonic velocity and compressive strength for high-volume mineral-admixtured concrete. *Cement and Concrete Research, 34*(12), 2329-2336.

Dukhin, A. S., Goetz, P. J., & Travers, B. (2005). Use of ultrasound for characterizing dairy products. *Journal of Dairy Science, 88*(4), 1320-1334.

Elvira, L., Rodríguez, J., & Lynnworth, L. C. (2009). Sound speed and density characterization of milk adulterated with melamine. *The Journal of the Acoustical Society of America, 125*(5), EL177-EL182.

Huppertz, T., & Kelly, A. L. (2006). Physical chemistry of milk fat globules. In Fox P. F. & McSweeney P. L. H. (Eds.) *Advanced Dairy Chemistry 2 Lipids*. Boston, MA: Springer.

Kyazze, G., & Starov, V. (2004). Viscosity of milk: Influence of cluster formation. *Colloid Journal, 66*(3), 316-321.

Langendorff, V., Cuvelier, G., Michon, C., Launay, B., Parker, A., & De kruif, C. G. (2000). Effects of carrageenan type on the behaviour of carrageenan/milk mixtures. *Food Hydrocolloids, 14*(4), 273-280.

Lee, S. C., Coan, B. S., & Bouxsein, M. L. (1997). Tibial ultrasound velocity measured in situ predicts the material properties of tibial cortical bone. *Bone, 21*(1), 119-125.

McClements, D. J. (1988). *The Use of Ultrasonics for Characterising Fats and Emulsions.* Doctoral dissertation, University of Leeds, UK.

McClements, D. J., & Gunasekaran, S. (1997). Ultrasonic characterization of foods and drinks: Principles, methods, and applications. *Critical Reviews in Food Science and Nutrition, 37*(1), 1-46.

5

McClements, D. J., & Povey, M. J. W. (1987). Ultrasonic velocity as a probe of emulsions and suspensions. *Advances in Colloid and Interface Science, 27*(3), 285-316.

Meyer, S., Rajendram, V. S., & Povey, M. J. W. (2006). Characterization of reconstituted milk powder by ultrasound spectroscopy. *Journal of Food Quality, 29*(4), 405-418.

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.

Ogasawara, H., Mizutani, K., Ohbuchi, T., & Nakamura, T. (2006). Acoustical experiment of yogurt fermentation process. *Ultrasonics, 44*, E727-E730.

Phillips, L. G., Mcgiff, M. L., Barbano, D. M., & Lawless, H. T. (1995). The influence of fat on the sensory properties, viscosity, and color of lowfat milk. *Journal of Dairy Science, 78*(6), 1258-1266.

Povey, M. J. W. (1989). Ultrasonics in food engineering part II: Applications. *Journal of Food Engineering, 9*(1), 1-20.

Sigfusson, H., Decker, E. A., & McClements, D. J. (2001). Ultrasonic characterization of Atlantic mackerel (Scomber scombrus). *Food Research International, 34*(1), 15-23.

Singh, A. P., McClements, D. J., & Marangoni, A. G. (2004). Solid fat content determination by ultrasonic velocimetry. *Food Research International, 37*(6), 545-555.

Walstra, P., Wouters, J. T., & Geurts, T. J. (2005). *Dairy Science and Technology*. Boca Raton FL: CRC press.

Wood, A. B. (1946). *Textbook of Sound*. London: George Bell and sons.

Yan, J., Wright, W. M. D., O'Mahony, J. A., Roos, Y., Cuijpers, E., & van Ruth, S. M. (2019). A sound approach: Exploring a rapid and non-destructive ultrasonic pulse echo system for vegetable oils characterization. *Food Research International, 125*, 108552.

Supplementary material

Table S5.1. Label information of the dairy products purchased from the local shops.

Sample	Brand	Fat (g/100 mL)	Carbohydrate (g/100 mL)	Protein (g/100 mL)	Minerals (g/100 mL)
Double cream 1	Avonmore	48	2.6	1.9	0.10
Double cream 2	Clonbawn	48	1.6	1.5	0.05
Double cream 3	Tesco	48	1.7	1.6	0.10
Normal cream 1	Clonbawn	38	3.0	2.0	0.10
Normal cream 2	Supervalu	40	3.1	2.0	0.20
Normal cream 3	Tesco	38	2.7	2.0	0.10
Light cream 1	Avonmore	26	3.3	2.4	0.10
Light cream 2	Clonbawn	25	4.3	2.8	0.08
Light cream 3	Supervalu	26	4.0	2.1	0.10
Whole milk 1	Avonmore	3.5	4.7	3.4	0.11
Whole milk 2	Supervalu	3.5	4.8	3.3	0.20
Whole milk 3	Tesco	3.5	4.8	3.5	0.10
Semi skimmed milk 1	Clonbawn	1.0	4.8	3.5	0.11
Semi skimmed milk 2	Supervalu	1.5	4.9	3.5	0.20
Semi skimmed milk 3	Tesco	1.5	4.9	3.5	0.10
Skimmed milk 1	Avonmore	0.2	5.2	3.8	0.12
Skimmed milk 2	Clonbawn	0.5	5.0	3.3	0.13
Skimmed milk 3	Tesco	0.3	4.8	3.5	0.10

5

Table S5.2. Means of ultrasonic velocity, viscosity [a] and density for the 18 commercial dairy samples and their diluted counterparts.

Sample [b]	Velocity (m/s)	Viscosity (mPa.s)	Density (g/L)	Sample [b]	Velocity (m/s)	Viscosity (mPa.s)	Density (g/L)
dbcr1	1491.4	36.6	974	wm1	1520.9	2.0	1036
dbcr1-0.8	1485.9	10.1	980	wm1-0.8	1520.9	1.9	1029
dbcr1-0.6	1488.1	5.3	985	wm1-0.6	1514.0	1.6	1022
dbcr1-0.4	1491.4	3.6	990	wm1-0.4	1507.2	1.6	1014
dbcr1-0.2	1492.5	3.1	995	wm1-0.2	1500.4	1.1	1007
dbcr2	1493.7	30.0	993	wm2	1519.8	2.2	1035
dbcr2-0.8	1485.9	7.7	994	wm2-0.8	1530.2	1.7	1028
dbcr2-0.6	1489.2	4.5	996	wm2-0.6	1520.9	1.5	1021
dbcr2-0.4	1492.5	3.2	997	wm2-0.4	1504.9	1.2	1014
dbcr2-0.2	1493.7	2.6	999	wm2-0.2	1499.3	1.1	1007
dbcr3	1483.7	17.0	981	wm3	1523.2	2.2	1037
dbcr3-0.8	1484.8	9.2	985	wm3-0.8	1511.7	1.8	1030
dbcr3-0.6	1489.2	4.8	989	wm3-0.6	1507.2	1.6	1022
dbcr3-0.4	1490.3	3.5	993	wm3-0.4	1502.6	1.3	1015
dbcr3-0.2	1492.5	2.2	996	wm3-0.2	1497.0	1.2	1007
cr1	1499.3	11.5	995	ssm1	1522.1	1.9	1037
cr1-0.8	1491.4	6.3	996	ssm1-0.8	1519.8	1.5	1030
cr1-0.6	1490.3	4.2	997	ssm1-0.6	1512.9	1.3	1022
cr1-0.4	1491.4	0.4	998	ssm1-0.4	1507.2	1.2	1015
cr1-0.2	1490.3	2.7	999	ssm1-0.2	1502.6	1.0	1007
cr2	1497.0	15.1	998	ssm2	1522.1	1.9	1037
cr2-0.8	1492.5	6.0	998	ssm2-0.8	1527.9	1.6	1030
cr2-0.6	1493.7	3.8	999	ssm2-0.6	1517.5	1.4	1022
cr2-0.4	1493.7	2.9	999	ssm2-0.4	1511.7	1.2	1015
cr2-0.2	1494.8	2.2	1000	ssm2-0.2	1501.5	1.1	1007
cr3	1492.5	11.8	980	ssm3	1525.6	1.8	1038
cr3-0.8	1492.5	6.9	984	ssm3-0.8	1514.0	1.6	1030
cr3-0.6	1492.5	3.9	988	ssm3-0.6	1506.0	1.5	1023
cr3-0.4	1494.8	3.2	992	ssm3-0.4	1502.6	1.2	1015
cr3-0.2	1493.7	2.2	996	ssm3-0.2	1498.1	1.0	1008
ltcr1	1507.2	86.4	1013	skm1	1525.6	2.3	1040
ltcr1-0.8	1503.8	0.0	1010	skm1-0.8	1522.1	2.0	1032
ltcr1-0.6	1500.4	0.0	1008	skm1-0.6	1515.2	1.7	1024
ltcr1-0.4	1497.0	0.0	1005	skm1-0.4	1508.3	1.5	1016
ltcr1-0.2	1495.9	0.0	1003	skm1-0.2	1501.5	1.1	1008
ltcr2	1503.8	6.1	1007	skm2	1523.2	1.7	1042
ltcr2-0.8	1499.3	0.0	1006	skm2-0.8	1519.8	1.5	1034
ltcr2-0.6	1498.1	0.0	1004	skm2-0.6	1512.9	1.3	1025
ltcr2-0.4	1497.0	0.0	1003	skm2-0.4	1507.2	1.2	1017
ltcr2-0.2	1497.0	0.0	1001	skm2-0.2	1501.5	1.1	1008
ltcr3	1508.3	28.6	1013	skm3	1526.7	1.7	1038
ltcr3-0.8	1502.6	0.0	1010	skm3-0.8	1525.6	1.5	1030
ltcr3-0.6	1500.4	0.0	1008	skm3-0.6	1516.3	1.4	1023
ltcr3-0.4	1495.9	0.0	1005	skm3-0.4	1507.2	1.3	1015
ltcr3-0.2	1494.8	0.0	1003	skm3-0.2	1501.5	1.1	1008

[a] Viscosity at shear rate 100 /s was used. [b] dbcr, double cream; cr, normal cream; ltcr, light cream; wm, whole milk; ssm, semi skimmed milk; skm, skimmed milk. The number after the letters represents the samples coding, last number after the hyphen represents the concentration of sample.

5
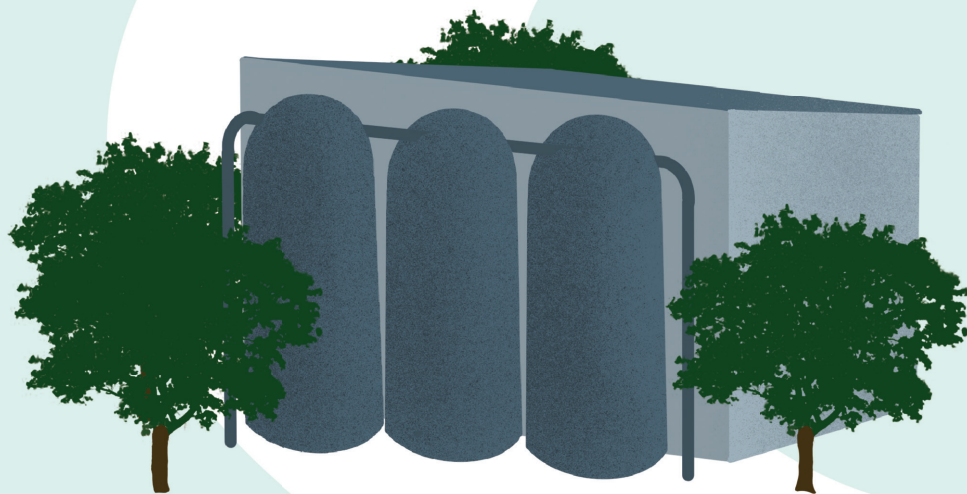
CHAPTER 6

Prevalence of milk fraud in the Chinese market and its relationship with fraud vulnerabilities in the chain

Abstract

This study aimed to assess the prevalence of ultra-high temperature (UHT) processed milk samples suspected of being adulterated on the Chinese market and, subsequently, relate their geographical origin to the earlier determined fraud vulnerability. A total of 52 UHT milk samples purchased from the Chinese market were measured to detect possible anomalies. The milk compositional features were determined by standardized Fourier transform-infrared spectroscopy, and the detection limits for common milk adulterations were investigated. The results showed that twelve of the analysed milk samples (23%) were suspected of having quality or fraud-related issues, while one sample of these was highly suspected of being adulterated (diluted with water). Proportionally, more suspected samples were determined among milks produced in the central-northern and eastern areas of China than in those from the north-western and north-eastern areas, while those from the south were in between. Combining the earlier collected results on fraud vulnerability in the Chinese milk chains, it appears that increased fraud prevalence relates to poorer business relationships and lack of adequate managerial controls. Since very few opportunities and motivations differ consistently across high and low-prevalence areas, primarily the improvement of control measures can help to mitigate food fraud in the Chinese milk supply chains.

Keywords

China; Fourier transform-infrared spectroscopy; Fraud vulnerability; Milk adulteration; Milk composition; One-class classifications

6.1. Introduction

The dairy industry in China has developed in parallel with the country's economic growth. The average annual milk consumption of Chinese consumers reached 36 kg per capita in 2016, though this is still lower than the average world consumption (Li, 2016). Liquid milk is the main type of dairy product in the Chinese market, with more than 90% of the raw milk being processed to liquid milk products (Ministry of Agriculture P. R. China, 2016). Moreover, the most popular liquid milk product, ultra-high-temperature (UHT) processed milk, accounted for more than half of the liquid milk consumption in 2018 (PwC, 2018).

The melamine infant formula incident in 2008 highlighted the vulnerability of the dairy industry in China and was a grave shock for this industry (Xiu & Klein, 2010). This incident resulted in great financial, as well as consumer confidence, losses for the sector. Milk, as a source of protein and calcium, plays an important role in the diet of Chinese consumers. Unfortunately, milk is one of the most commonly adulterated foods due to its popularity, production and sales in mass volume levels and the price paid for the product based on its composition (Moore, Spink, & Lipp, 2012). Moreover, the number associated with the adulteration of dairy products made it the top-ranked product among the animal food products that are typically subjected to fraud (Zhang & Xue, 2016). The rapidly growing economy in China has led to a gradual shift of concerns about food security to food safety (Lam, Remais, Fung, Xu, & Sun, 2013). Linked to food safety is also the prevalence of food fraud, as the adulteration of milk products, depending on the type of adulterant used, can have safety implications for a food product.

Food fraud has become a widely acknowledged concern, not only within the food supply chain, but also more widespread, for instance, by consumers. Recent studies on Chinese consumers' preferences showed that Chinese consumers are concerned about the risk of domestic milk products and, especially, infant milk formula (Li, Sijtsema, Kornelis, Liu, & Li, 2019) and therefore prefer to buy "foreign milk powder" instead of domestic products (Yin, Li, Xu, Chen, & Wang, 2017). Melamine is not the only milk adulterant. Milk adulterations range from very simple, such as dilution, to very complex, such as synthesizing milk with urea, vegetable oil, detergents and other chemical compounds (Handford, Campbell, & Elliott, 2016). A number of substances have been listed as potential milk adulterants by different scholars (Handford et al., 2016; Hansen & Holroyd, 2019).

In our previous studies, technical opportunities were identified as medium-high risk factors in the investigations of the fraud vulnerability of milk supply chains, whereas the detectability of milk adulteration was a main concern of the participants in the supply chain (Yang et al., 2019; Yang, Huisman, Hettinga, Zhang, & van Ruth, 2020). Various advanced techniques have been developed for milk authentication. For example, chromatographic methods combined with mass spectrometry (MS) have been used for detecting milk adulteration with nitrogen-rich compounds (Abernethy & Higgs, 2013; Tittlemier, 2010) and vegetable oils (Nurseitova et al., 2019), digital imaging for milk protein determinations (Silva & Rocha, 2020)

and the detection of hydrogen peroxide in milk (Lima, Rossini, Pezza, & Pezza, 2020), proton transfer reaction mass spectroscopy (PTR-MS) and stable isotope ratio mass spectroscopy (IRMS) for the discrimination of organic milk (Chung, Park, Yoon, Yang, & Kim, 2014; Liu, Koot, Hettinga, de Jong, & van Ruth, 2018) and nuclear magnetic resonance (NMR) for the nontargeted detection of multiple adulterants in milk powder (Bergana, Adams, Harnly, Moore, & Xie, 2019). Moreover, infrared spectroscopy-based techniques have become the most commonly used method for determining food authenticity, and they are considered as alternatives to reference methods (Kamal & Karoui, 2015). Furthermore, automated equipment based on Fourier transform-infrared (FTIR) spectroscopy has been developed to determine milk composition, providing high analytical capacity and low operational costs (Sánchez et al., 2007). FTIR spectroscopy is extensively used worldwide for milk quality control, because little sample preparation is needed, and the analysis is rapid. In combination with statistical analysis, FTIR spectroscopy has been applied to identify several milk adulterants, such as melamine (Jawaid, Talpur, Sherazi, Nizamani, & Khaskheli, 2013), whey protein (Cassoli, Sartori, Zampar, & Machado, 2011), sodium bicarbonate, sodium citrate, and corn starch (Coitinho et al., 2017).

Considering the previous milk fraud incidents in China and the uprising public concerns from the Chinese consumers, there is an urgent demand for information on the integrity (i.e., safety, quality and authenticity) of milk products on the Chinese market to restore consumers' trust. As previous studies have already addressed food safety issues (Du et al., 2019; Li et al., 2017; Xiong et al., 2020; Zheng et al., 2013), the current study aims to evaluate the occurrence of milk adulteration in China and relate the prevalence of the milk fraud to fraud vulnerability profiles. The occurrence of milk adulteration was evaluated by anomaly detection. Anomaly detection was firstly conducted based on individual measurement parameters that are commonly used in practice, using both univariate and multivariate approaches. For both approaches, the data of a control group of genuine UHT milk samples was used in combination with groups of protein-rich, nitrogen-based and carbohydrate-based milk adulterants, as well as nonallowed preservatives at various concentration levels to set boundaries. Finally, commercial UHT milk samples from different regions in China were tested against the boundaries using both developed approaches. Fraud prevalence was related to the geographical origin of the samples and compared with previously established fraud vulnerabilities for the geographical areas.

6.2. Materials and methods

6.2.1. Sample collection

A total of twelve UHT milk samples from five different brands were purchased from local supermarkets in Beijing in November 2018. They were from the top 10 dairy-processing enterprises in China and considered as the reference samples. Among these twelve samples, four samples were labelled as premium quality (protein content > 3.5%) and produced in the north of China, four were of normal quality and also produced in the north and the remaining four samples were of normal quality and produced in the south. Accordingly, three milk pools

were prepared: the first milk pool (Pool A) was a mixture of the four samples of premium quality, the second pool (Pool B) a mixture of the four samples of normal quality from the north of China and the third pool (Pool C) a mixture of the four samples of normal quality from the south of China. The ratio of the four milk samples in each pool was 1:1:1:1 w/w. Three milk pools of 100 g each were prepared, to which an adulterant was added at several concentrations. The twelve reference samples and the three milk pools were from major producers and, hence, considered as the control samples (i.e., nonadulterated). These 15 samples comprised the training set.

A total of 52 commercial UHT milk samples were purchased from local markets (in Beijing) and e-commerce (across China) during the winter of 2018/2019 (December 2018 to January 2019), and these samples comprised the market survey test set. The distribution of the geographical origin of the market survey samples is shown in Fig. 6.1.
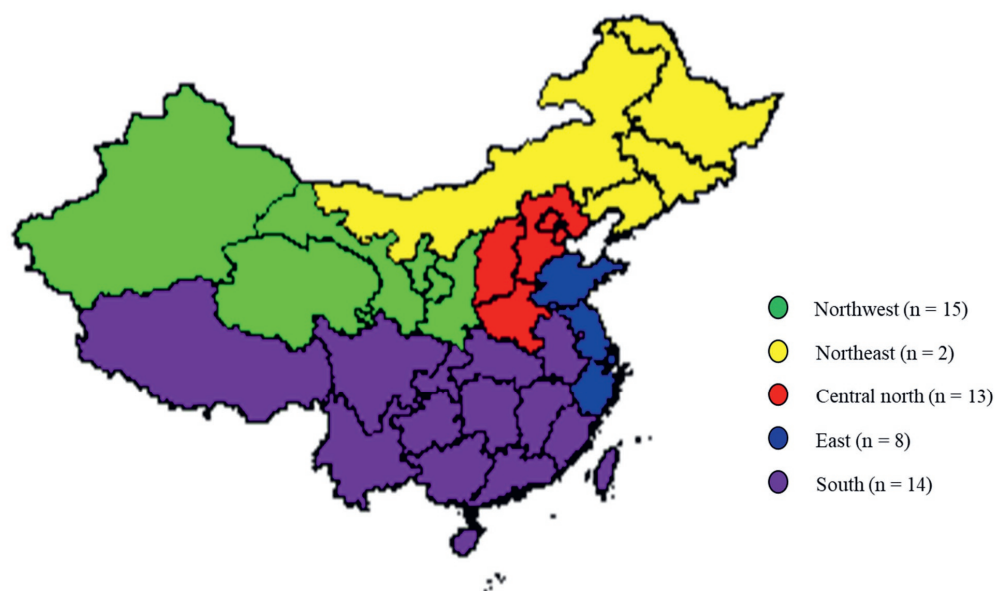

Fig. 6.1. The geographical distribution of the market survey samples.

6.2.2. Adulterations and measurements

Following the same procedure of a comparable study that was conducted on milk samples from the Dutch market, the same adulterants and the same adulteration levels were applied (Yang, Hettinga, Erasmus, Pustjens, & van Ruth, 2020). A total of 24 adulterants were used, which were categorized into five groups 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) and (5) water. Both single and combined adulterations were conducted. The single adulterations were carried out at four levels for each adulterant. The formulas and detailed plan for the single adulterations are provided in Table S6.1 (Supplementary material). The combined adulterations were conducted in two steps: first, 40 g water was added to 100 g of a milk pool sample; then, one of the adulterants from either the protein-rich, nitrogen-based or carbohydrate-based adulterant category was added to the diluted milk 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 total solids (TS)/milk TS content for the carbohydrate-based adulterations). The detailed information of the combined adulteration is provided in Table S6.2 (Supplementary material). Ultimately, a total of 288 single-adulterated samples and 51 combined-adulterated samples were prepared. These 339 adulterated samples were considered as the adulterant test set.

MilkoScan FT120 equipment (Foss Electric, Hilleroed, Denmark) was used to measure the milk composition. The equipment is based on the FTIR technique and reports a series of milk compositional parameters, namely protein, fat, lactose, total solids (TS), solids non-fat (SNF) content, density and freezing point depression (FPD). The FTIR spectra were not acquired separately because of the limitation of the instrument used. All the samples were prepared at room temperature and measured within two hours after preparation.

6.2.3. Statistical analysis

6.2.3.1. Univariate analysis: determination of boundaries for each variable

The mean and standard deviation (SD) for the seven variables (i.e., protein, fat, lactose, total solids, solids non-fat, freezing point depression and density) were calculated based on the 15 control samples. Next to that, the values of the 0.5^{th} and 99.5^{th} percentiles for each variable were used to set boundaries for anomaly detection.

According to a programme of measurements of over 3 million raw milk samples for legislatorial control (Zuivelverordening, 2000), the standard deviation of this large sample set for the seven variables is roughly double the values of the control samples in this study. To adapt the variance of the control samples in a practical way, the data of the 15 control samples was transformed into a variance-adjusted dataset, where the mean value for each variable remained the same, but the SD value was adjusted to twice the measured SD. The new dataset was converted from the measured data for each variable separately using Eq. (6.1):

$$X_{new} = \left(\frac{X-\mu}{\sigma} \times 2\sigma\right) + \mu \qquad\qquad \text{Eq. (6.1)}$$

where X_{new} is the variance-adjusted data, X is the measured data for the control samples, μ is the mean value of the 15 control samples and σ is the SD of the 15 control samples. Next, the variance-adjusted boundaries were determined using the values of 0.5th percentile and 99.5th percentile for the seven variance-adjusted variables.

The results of the measured and variance-adjusted datasets are shown in Tables S6.3 and S6.4 (Supplementary material), respectively. Both the measured boundaries and variance-adjusted boundaries were then utilized for both the adulterant test set and market survey test set. The univariate calculations were performed using Microsoft Excel 2016 (Microsoft, Redmond, WA, US).

6.2.3.2. Multivariate analysis: determination of boundaries for milk with one-class classification models

One-class classification (OCC), which focuses on a single target class, has become a common modelling approach for the verification of food authenticity (Oliveri, 2017). Three one-class classification (OCC) models were applied in this study, namely k-nearest neighbours (KNN), soft independent modelling of class analogies (SIMCA) and support vector machine (SVM). KNN has no requirement for the data distribution and is robust to noisy training data, and hence, it is suitable for analysing small training sets (Beebe, Pell, & Seasholtz, 1998). SIMCA focuses more on the similarities among samples within a class and is thus widely used for OCC models (Gurbanov, Gozen, & Severcan, 2018). SVM is another fitting approach that can be applied to datasets with a limited number of training samples (Gholami & Fakhari, 2017). SVM evaluates the distance from an object to the boundary. For this study, the Gaussian radial basis function (RBF) kernel was used to determine the boundary for the SVM model.

The measured dataset ($n = 15$) and variance-adjusted dataset ($n = 15$) of the control samples were separately used as the training set for the model development. The training set was subjected to leave-30%-out cross-validation with 100 repetitions. The dataset was pre-processed by means of autoscaling. Next to that, the three classifiers (KNN, SIMCA and SVM) were applied. A significant level of 1% ($p < 0.01$) was used for determining the critical classification thresholds. The adulterant test set, comprising 339 adulterated milk samples, was then subjected to the developed models. The three OCC models were evaluated applying the following parameters: the k value for the KNN model was selected from consecutive numbers 1–10; the number of factors n for SIMCA was selected from consecutive numbers 1-7; γ in the Gaussian radial basis function (RBF) kernel 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 average value of the percentages of correctly assigned samples for the cross-validation set and adulterant test set was used to evaluate the overall performance of the models. The optimal parameter for the best performing model was determined accordingly. Next, the market survey test set, comprising 52 samples, was subjected to the selected models. The OCC model development in this study was performed using R 3.6.1 (R Foundation for Statistical Computing, Vienna, Austria).

6

6.2.3.3. Exploratory analysis and regression model

The result of the compositional features was subjected to principal component analysis (PCA) to visualize the grouping of the control samples and market survey samples after pre-processing by autoscaling. The PCA was performed by R 3.6.1 (R Foundation for Statistical Computing, Vienna, Austria). Principal component regression (PCR) with leave-one-out cross-validation was conducted between the geographical prevalence of the suspected samples and the result of the food fraud vulnerability assessment in the corresponding areas. The mean ranks of the scores of fraud factors for four of the main milk production areas in China (i.e., central-north, northeast, northwest and east of China) from a previous fraud vulnerability assessment study (Yang et al., 2020) were used to develop a model to predict the percentage of suspected samples in these areas after pre-processing by autoscaling. The PCR was performed using Pirouette 4.5 (Infometrix Inc., Bothell, WA, USA).

6.3. Results and discussion

6.3.1. Control samples

6.3.1.1. Natural variation of the control samples

A certain degree of variation in the milk composition was observed among the three milk pools analysed by FTIR spectroscopy (Table 6.1). All the measured compositional features of the premium milk pool (pool A) were higher than those of the normal milk pools from both the north and the south (pool B and C). Generally, the fat and lactose contents of the control milk are in agreement with those of the raw milk in China, which are 3.6% to 4.2% w/w and 4.7% to 5.1% w/w, respectively (Guo, Liu, Xu, & Xia, 2010; Yang, Yang, Yi, Pang, & Xiong, 2013). However, the protein content of the control samples (3.4% to 3.7% w/w) was higher than that of the raw milk produced by Chinese Holstein cattle, which is approximately 2.9% to 3.3% w/w (Li, 2016; Yang et al., 2013). This difference may have been caused by protein standardization techniques used during the processing of the milk, for instance, flash evaporation (Zhao, Yue, Wang, & Peng, 2005). It may also be due to the use of raw milk of a higher protein content from other dairy cattle breeds. A difference in milk composition was observed between the commercial UHT milk samples from the Dutch and Chinese markets. The means of the fat content and lactose content of the Chinese samples (4.0% and 5.1% w/w, respectively) were slightly higher than those of the Dutch ones (3.8% and 4.7% w/w, respectively), while the mean values for the protein content of the milk from the two countries were more or less the same (3.5%-3.6% w/w) (Yang, Hettinga, Erasmus, Pustjens & van Ruth, 2020).

Table 6.1. Means and standard deviation (SD) of the compositional features of the control samples measured by standardised Fourier transform-infrared spectroscopy and the boundaries based on the measured dataset and variance-adjusted dataset.

Dataset		Compositional features [a]						
		Protein (% w/w)	Fat (% w/w)	TS (% w/w)	SNF (% w/w)	Lactose (% w/w)	Density (g/L)	FPD (°C)
Pools	Pool A (premium, north)	3.69	4.05	13.72	9.73	5.30	1034	0.567
	Pool B (normal, north)	3.44	3.75	12.79	9.05	4.88	1031	0.524
	Pool C (normal, south)	3.47	3.83	13.04	9.24	5.04	1032	0.544
Measured dataset	Mean	3.54	3.95	13.24	9.33	5.06	1032	0.543
	SD	0.15	0.24	0.51	0.35	0.21	1	0.022
Measured boundary	Lower boundary	3.33	3.60	12.57	8.94	4.80	1031	0.516
	Upper boundary	3.73	4.42	14.03	9.85	5.39	1035	0.576
Variance-adjusted boundary	Lower boundary	3.13	3.26	11.90	8.55	4.54	1030	0.489
	Upper boundary	3.93	4.90	14.82	10.37	5.72	1038	0.608

[a] TS, total solids; SNF, solids non-fat and FPD, freezing point depression.

6.3.1.2. Control samples and univariate detection approach

The univariate boundaries based on the measured dataset and variance-adjusted dataset are presented in Table 6.1. The 15 control samples were tested by the two sets of boundaries. According to the measured dataset, five compositional values of the control samples out of 105 measurements exceeded the measured boundaries, including four that exceeded the upper boundaries and one that exceeded the lower boundary. The samples exceeding the upper boundaries concerned protein, fat, SNF and the lactose content (control samples 2-4), while control sample 8 exceeded the lower boundary of the TS content (Table S6.3, Supplementary material). It seems that the samples of which the compositional features exceeded the upper boundaries are due to features generally occurring in premium quality milk; in other words, they are in the top 0.5% of the distribution of the control samples with regard to the protein, fat, SNF and/or lactose concentration. Considering the lactose content of raw milk is quite stable, ranging from 4.5% to 5.0% w/w (Fox, McSweeney, & Paul, 1998; Heck, van Valenberg, Dijkstra, & van Hooijdonk, 2009), and would not be intentionally adjusted during the milk-processing, it is believed that the high lactose content of the premium milk product was caused by the use of flash evaporation, which is sometimes used for the production of premium milk in China (Zhao et al., 2005). Such flash evaporation would also lead to an increase of the protein, fat and SNF contents by water removal, as was found for samples 2-4. As samples 2-4 were thus exceeding the boundaries for reasons other than adulterations, they were kept in the control group. Considering the large variance that would be found in practice among unadulterated

6

samples, the variance-adjusted dataset was also applied, against which all the 15 controls samples were considered normal.

6.3.1.3. Control samples and the multivariate detection approach

One-class classification models using three classifiers (KNN, SIMCA and SVM) were calculated, and for each classifier, the model with the best performance was selected. The results of the selected models are shown in Table 6.2. For all the three classifiers in both scenarios (i.e., the measured dataset and variance-adjusted dataset), the samples of the training set were 100% correctly classified. The KNN classifier performed a bit better for the cross-validation set, achieving 92% and 93% accuracy for the model of the measured dataset and variance-adjusted dataset, respectively. Combining the performance of the models in both scenarios, KNN was selected as the best classifier for the OCC model for further analysis.

Table 6.2. The results of the one-class classification models developed from the measured dataset and variance-adjusted dataset. All the present values refer to the average of 100 repetitions of cross-validation for the corresponding dataset.

Model	Performance for dataset	Correctly assigned samples [a] (%)		
		KNN	SIMCA	SVM
Model developed from the measured dataset	Training set	100	100	100
	Cross validation set	92	88	90
	Adulterant test set	77	75	79
	Overall performance	84	81	84
Model developed from the variance-adjusted dataset	Training set	100	100	100
	Cross validation set	93	91	92
	Adulterant test set	66	60	63
	Overall performance	79	75	77

[a] KNN stands for k-nearest neighbours; the KNN model with the best performance was estimated with $k = 3$ for both the measured and variance-adjusted datasets. SIMCA stands for soft independent modelling of class analogies; the SIMCA model with the best performance was estimated with the number of the factors $n = 3$ for both the measured and variance-adjusted datasets. SVM stands for support vector machine; the SVM model with the best performance was estimated with $\gamma = 0.1$ for both the measured and variance-adjusted datasets.

To summarise, based on the variation of the composition of both the control samples in this study (i.e., the measured dataset) and a more practicable scenario (i.e., the variance-adjusted dataset), the univariate boundaries and multivariate models were determined, respectively. The same models were then subjected to the adulterant test set and market survey set.

6.3.2. Adulterants

6.3.2.1. Adulterants and the univariate detection approach

To test the detection capacity of the developed approaches, the univariate boundaries were first applied to the adulterant test set. As expected, the univariate boundaries of the measured dataset flagged more adulterations than the variance-adjusted boundaries, as shown in Fig. 6.2.

Both boundaries flagged the high concentrations (levels 2-4) of the protein-rich adulterations and almost all carbohydrate adulterations (except for starch adulteration). Furthermore, the measured boundaries flagged high concentrations (levels 3-4) of the nitrogen adulterants, while the variance-adjusted boundaries had a lower performance for these adulterations. The water dilutions were almost universally flagged by both boundaries. Most of the preservative concealers passed unnoticed for both types of boundaries (Fig. 6.2).

Fig. 6.2. The results of the adulterant test sets for the three milk pools based on (a) the measured boundaries of the univariate detection and (b) the variance-adjusted boundaries of the univariate detection, indicating the potential to identify suspected milk adulterations. The samples with all results within the boundaries are coloured green, while the rest is coloured red. The full names of the adulterants are shown in the abbreviations list. N.A., not applicable.

6.3.2.2. Adulterants and multivariate detection approach

The selected KNN models were also applied to the adulterant test set. Similar to the scenario of the univariate detection, the KNN model based on the measured dataset also raised more flags than that based on the variance-adjusted dataset for the various adulterations, as shown in Fig. 6.3. All carbohydrates, except for the starch adulterations, were flagged by both models. When considering the performance of the KNN model of the measured dataset, the average specificities of the model of the 100 repetitions for the protein-rich adulterations (95%) and carbohydrate adulterations (93%) were slightly higher than that of the nitrogen adulterations (86%), only the ammonia sulphate adulterations in the latter group were fully

flagged. The KNN model of the measured dataset flagged most water dilutions, while that of the variance-adjusted dataset flagged no water dilutions at all. The specificities of both models for the preservative adulterations were very low.

Fig. 6.3. The results of the adulterant test sets for the three milk pools based on the threshold of the of k-nearest neighbour (KNN) 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 adulterants are shown in the abbreviations list. N.A., not applicable.

6.3.2.3. Comparison of approaches

For both the univariate and multivariate approaches, the ones based on the measured dataset flagged more adulterated samples than those based on the variance-adjusted dataset, due to the measured dataset having less variance. All the developed criteria succeeded to flag carbohydrates, except for the starch adulterations, and also flagged protein-rich adulterations at the higher levels. Both approaches based on the measured dataset could identify nitrogen adulterations at higher levels as well. In addition, it is noted that the multivariate approach did not perform better than the univariate approach in distinguishing the milk adulterations.

6.3.3. Market survey samples: What type of suspected milk samples are discovered using the developed approaches?

6.3.3.1. Suspected samples flagged by the univariate detection approach

As the developed approaches showed different abilities in detecting adulterations, they were all applied to the market survey samples. Out of the 52 samples from the market survey, 37 samples were flagged according to the univariate boundary of the measured dataset (Table 6.3). Their compositional results showed that the protein, fat and lactose contents were the main parameters that exceeded the boundaries. Only two samples were flagged for exceeding the upper boundary based on the lactose content (samples 32 and 37). As discussed in Section 6.3.1.2, the lactose content of raw milk is rather stable and will be below 5.0% w/w. The extremely high lactose content of these two suspected samples (> 5.4% w/w) was likely caused by some kind of manipulation during processing, by either legal ways like flash evaporation or illegal ways like carbohydrate or dairy powder additions. The other 35 suspected samples were flagged, because their compositional variables exceeded the lower boundaries. Among these suspected samples, 32 samples were observed to be deficient in proteins, 15 samples deficient in fat and 15 samples deficient in lactose. It is not surprising that the TS and SNF contents of these suspected samples exceeded the lower boundary as well. These results indicate that the flagged samples were deviating from the control samples with respect to multiple compositional parameters, including protein, fat and/or lactose contents. However, it should be noted that the variance of the measured dataset is smaller than that faced in practice, which likely resulted in more genuine samples being misclassified. This would hence increase the workload of further checking these samples and lower the users' acceptance. In addition, although the protein or fat contents of these samples were lower than the boundary, they were not exceeding the lower limit of the national food safety standard for sterilized milk, where 2.9% w/w is stipulated for the protein content and 3.1% w/w for the fat content (National Standard of the P. R. China, 2010). It seems that the boundaries based on the measured dataset were thus too strict for practical use.

A total of twelve samples were flagged according to the univariate boundary based on the variance-adjusted dataset (Table 6.3). Eleven of these flagged samples were lower in protein contents, exceeding the lower boundary (3.13% w/w) of the variance-adjusted dataset. In addition, it was observed that some of the other compositional features such as the fat, total solids, density or FPD of these samples also exceeded the respective lower boundaries. There are probably multiple reasons for the low protein contents in these UHT milk products. One is that the raw milk used to produce the final products could have been low in protein contents. Considering there is no prohibition on adjusting the protein, fat, or lactose contents of UHT milk in China, another reason may be that the milk composition was changed during the processing—for instance, removing part of the milk fat would result in a lower fat content. It is also possible that the milk samples were diluted, although this would be a violation of the national food safety standard for sterilized milk, if the milk was diluted during the processing but labelled as "pure milk" on the package of the final product (National Standard of the P. R.

China, 2010). Since the milk samples in this study were all labelled "pure milk", the lower protein and fat contents of the flagged milk samples might be indicators of potential milk manipulation. Therefore, these twelve samples were suspected of quality or fraud issues. In addition, the FPD of one suspected sample (sample 1) exceeded the boundary of the variance-adjusted dataset. The freezing point is principally affected by the lactose and dissolved salts in the milk and is very constant due to its effect on the osmotic pressure of milk. The addition of water would reduce the concentration of these compounds and lead to a change of the freezing point towards zero (De Longhi et al., 2012). Taking the reasons above into consideration, sample 1 was most likely adulterated by water dilution.

6.3.3.2. Suspected samples flagged by the multivariate detection approach

According to the KNN OCC model developed with the measured dataset, 29 market survey samples were flagged (Table 6.3). Among these suspected samples, the compositional features of 24 samples (samples 1-16, 18-22 and 26-28) were in-line with the univariate analysis, which exceeded the lower boundary based on the measured dataset. The five other samples (samples 38-42), of which the compositional features were within the univariate boundaries, could be explained by the characteristics of the KNN algorithm. In the classification phase of the KNN OCC model, for each object in the test set, the k ($k = 3$ in this study) nearest training set vectors (the control samples) are determined, the distance between them calculated and the classification is then done by comparing the distance between the object and its k nearest neighbours to a predetermined threshold (Schliep, Hechenbichler, & Lizee, 2016). To visualise the variation among samples, a PCA was performed with the measured data of the control samples and market survey samples, as presented in Fig. 6.4. Although the samples 38-42 were located in the middle of the control samples, their k nearest neighbours were not as close as the nonflagged samples, explaining why they were flagged by the KNN OCC model. It is noted that Fig. 6.4 shows only the first two PCs of the PCA, instead of the "complete distribution", as used for the KNN model; however, it does provide a visualisation of the variation among samples.

Three samples (samples 38, 39, and 43) were flagged by the KNN OCC method developed from the variance-adjusted dataset (Table 6.3). Similar to the scenario as described for the measured dataset, the spatial distance between these three flagged samples and their k ($k = 3$ in this study) nearest neighbours is larger than the determined threshold of the models based on the variance-adjusted dataset. As a result, these samples were flagged as differing from the control group.

Table 6.3. The results of the suspected samples in the market survey set according to the developed detections [a].

ID	Protein (% w/w)	Fat (% w/w)	TS (% w/w)	SNF (% w/w)	Lactose (% w/w)	Density (g/L)	FPD (°C)	Area	Province	Univariate boundaries		Multivariate models (KNN)	
										Measured dataset	Variance-adjusted dataset	Measured dataset	Variance-adjusted dataset
1	3.09 **	3.56*	12.06 *	8.42 **	4.58 *	1029 **	0.488 **	N	Tianjin	7	4	yes	no
2	3.07 **	3.48*	11.94 *	8.37 **	4.55 *	1029 **	0.490 *	NW	Xinjiang	7	3	yes	no
3	3.28 *	3.49*	12.25 *	8.69 *	4.65 *	1030 *	0.515 *	NW	Xinjiang	7	0	yes	no
4	3.05 **	3.79	12.35 *	8.51 **	4.71 *	1029 **	0.498 *	N	Henan	6	3	yes	no
5	3.12 **	3.61	12.18 *	8.50 **	4.62*	1030 *	0.492 *	N	Tianjin	6	2	yes	no
6	2.99 **	3.43*	12.06 *	8.58 *	4.85	1030 *	0.512 *	E	Zhejiang	6	1	yes	no
7	3.14 *	3.70	12.32 *	8.56 *	4.67 *	1030 *	0.494 *	N	Henan	6	0	yes	no
8	3.28 *	3.74	12.49 *	8.68 *	4.64 *	1030 *	0.513 *	NW	Xinjiang	6	0	yes	no
9	3.17 *	3.87	12.54*	8.62 *	4.69 *	1030 *	0.505 *	NW	Xinjiang	6	0	yes	no
10	3.22 *	3.40 *	12.19 *	8.72 *	4.74 *	1031	0.497 *	S	Chongqing	6	0	yes	no
11	3.33	3.60 *	12.39 *	8.73 *	4.63 *	1030 *	0.492 *	S	Yunnan	6	0	yes	no
12	2.95 *	3.05 **	11.78 **	8.69 *	4.99	1031	0.510 *	E	Jiangsu	5	3	yes	no
13	3.14 *	4.16	12.75	8.55 *	4.65 *	1029 **	0.496 *	N	Tianjin	5	1	yes	no
14	3.03 **	3.57 *	12.34 *	8.74 *	4.96	1031	0.513 *	S	Yunnan	5	0	yes	no
15	3.17 *	3.40 *	12.29 *	8.85 *	4.93	1031	0.510 *	N	Shanxi	5	0	yes	no
16	3.15 *	4.15	12.81	8.62 *	4.70 *	1030 *	0.500 *	NW	Xinjiang	5	0	yes	no
17	3.27 *	3.53 *	12.41 *	8.83 *	4.80	1031	0.514 *	S	Yunnan	5	0	no	no
18	3.01 **	3.63	12.26 *	8.58 *	4.81	1028 **	0.538	S	Hubei	4	1	yes	no
19	3.12 **	3.51 *	12.47 *	8.95	5.07	1032	0.513 *	S	Yunnan	4	1	yes	no
20	3.25 *	4.21	13.05	8.81 *	4.80	1030 *	0.505 *	N	Hebei	4	0	yes	no
21	3.12 **	4.06	12.86	8.78 *	4.91	1030 *	0.523	NW	Shaanxi	3	1	yes	no
22	3.25 *	3.31 *	12.32 *	8.97	4.96	1032	0.520	N	Beijing	3	0	yes	no
23	3.37	4.16	13.09	8.89 *	4.75 *	1030 *	0.525	NW	Gansu	3	0	no	no
24	3.30 *	3.47 *	12.50 *	8.99	4.94	1032	0.523	S	Guangdong	3	0	no	no
25	3.26 *	3.58 *	12.67	9.07	5.06	1032	0.531	N	Hebei	2	0	no	no
26	3.09 **	3.75	12.76	9.00	5.16	1032	0.541	E	Shandong	1	1	yes	no
27	3.26 *	4.23	13.24	9.00	4.97	1031	0.570	NW	Qinghai	1	0	yes	no
28	3.32 *	3.97	13.24	9.29	5.21	1033	0.536	S	Yunnan	1	0	yes	no
29	3.29 *	3.84	12.91	9.05	5.00	1032	0.538	E	Shandong	1	0	no	no

6

ID	Protein (% w/w)	Fat (% w/w)	TS (% w/w)	SNF (% w/w)	Lactose (% w/w)	Density (g/L)	FPD (°C)	Area	Province	Univariate boundaries Measured dataset	Univariate boundaries Variance-adjusted dataset	Multivariate models (KNN) Measured dataset	Multivariate models (KNN) Variance-adjusted dataset
30	3.29 *	3.83	12.94	9.06	5.05	1032	0.532	N	Beijing	1	0	no	no
31	3.30 *	3.91	12.95	9.03	4.97	1032	0.524	N	Hebei	1	0	no	no
32	3.63	3.88	13.67	9.83	5.42†	1035	0.573	NE	Heilongjiang	1	0	no	no
33	3.34	3.57 *	12.73	9.13	5.03	1032	0.524	NW	Ningxia	1	0	no	no
34	3.26 *	3.99	13.10	9.11	5.08	1032	0.539	NW	Xinjiang	1	0	no	no
35	3.32 *	3.97	13.17	9.19	5.11	1032	0.538	S	Chongqing	1	0	no	no
36	3.28 *	3.63	12.67	9.01	4.97	1032	0.530	S	Guangdong	1	0	no	no
37	3.58	3.85	13.57	9.74	5.40†	1035	0.566	S	Guangdong	1	0	no	no
38	3.47	4.32	13.73	9.44	5.21	1033	0.546	E	Jiangsu	0	0	yes	yes
39	3.60	3.86	13.40	9.55	5.18	1033	0.553	NE	Heilongjiang	0	0	yes	yes
40	3.55	4.20	13.55	9.35	5.02	1032	0.540	E	Shandong	0	0	yes	no
41	3.40	3.72	13.08	9.36	5.20	1033	0.548	NW	Xinjiang	0	0	yes	no
42	3.40	4.40	13.56	9.16	4.99	1032	0.536	S	Guizhou	0	0	yes	no
43	3.51	4.22	13.83	9.65	5.36	1034	0.560	N	Hebei	0	0	no	yes

[a] ID, sample identification; TS, total solids; SNF, solids non-fat; FPD, freezing point depression; KNN, k-nearest neighbours. S, south; E, east; NW, northwest; NE, northeast and N, central-north. The colours dark-red, light-red and green indicate the worst-to-no anomalies detected for the univariate boundaries, and the colours red and green indicate the abnormal samples marked by the KNN models. * The value exceeding the lower measured boundary. ** The value exceeding the lower variance-adjusted boundary. † Value exceeding the upper measured boundary.
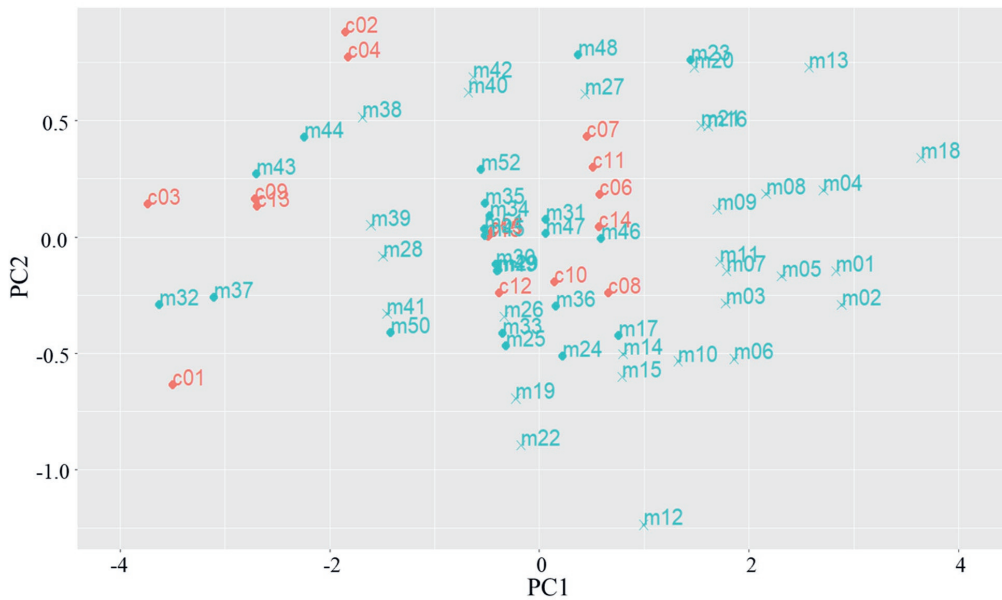
Fig. 6.4. The principal component analysis (PCA) plot of the first two PC dimensions of the control samples (red points) and the market survey samples (blue points) based on the results obtained from the MilkoScan measurements. The suspected samples flagged by the k-nearest neighbours (KNN) model of the measured dataset are cross-shaped (x).

6.3.3.3. Overall suspected samples of the market survey set

A total of 43 samples of the market survey set (n = 52) were flagged by the developed approaches (Table 6.3). Some exceeded only the boundary of one feature, while in most cases, samples exceeded multiple boundaries. Since more variation, as would be expected in practice, has been considered with the application of the variance-adjusted dataset, more attention should be paid to the samples exceeding the variance-adjusted boundaries. As every approach has its limits, combining multiple criteria simultaneously to detect suspected samples would provide a new perspective. In the end, the samples that violated most criteria, i.e., three out of four, were considered as the suspected samples among the market survey samples, which resulted in twelve suspected samples (samples 1-2, 4-6, 12-15, 18, 19, 21 and 26). Among these samples, four were produced in the central-northern area of China, three in the eastern area, two in the north-western area and three in the southern area.

6.3.4. Relation between the origin of the suspected milk and the previously determined fraud vulnerability

As shown in Table 6.4, the percentages of the suspected samples in the north-eastern and north-western areas (0% and 13%, respectively) were lower than those in the central-northern and eastern areas (31% and 38%, respectively). This is in-line with a study on the regional distribution of reported food fraud incidents, where more food fraud scandals or incidents were reported in the provinces of the central-northern and eastern areas compared to the north-

western and north-eastern areas in China (Zhang & Xue, 2016). The fraud vulnerabilities of the northern and eastern areas have been identified in a previous study (Yang et al., 2020). The fraud factors showing significant differences between the milk production areas are presented in Table 6.4, following the result of a PCR model that aimed to relate the determined fraud vulnerability to the number of suspected samples in certain areas.

Table 6.4. The geographical prevalence of the suspected samples (flagged by the three criteria in Table 6.3), the mean ranks of scores of the fraud factors showing significant differences between the four geographical areas [a] and the principal component regression (PCR) results.

Parameters		East	Central-north	North-west	North-east	Coeffi-cients [d]
Percentage (%) of suspected samples in the market survey set (number of suspected/total samples)		38% (3/8)	31% (4/13)	13% (2/15)	0% (0/2)	-
Fraud factors on opportunities and motivations [b]	1. Available technology for milk adulteration	49	46	71	66	-0.201
	2. Detectability of adulteration	51	56	35	58	-0.056
	3. Accessibility to production activities	53	49	57	63	-0.243
	4. Relationships within the supply chain	47	61	39	36	0.174
	5. Valuable components/attributes	38	62	39	42	0.023
	6. Farmer's financial pressure imposed by the company	40	61	40	45	0.005
	7. Level of competition	73	41	73	53	0.096
	8. Price difference due to regulatory differences	62	47	67	48	0.125
Fraud factors on Controls [c]	9. Application of integrity screening of employees in the company	51	46	70	63	-0.172
	10. Strictness of the ethical code of conduct in the company	45	48	70	61	-0.188
	11. Support of a whistle-blowing system in the company	62	47	69	48	0.115
	12. Specificity of the national food policy	70	48	55	49	0.209
	13. Availability of a fraud contingency plan	62	45	65	61	-0.051

[a] The fraud vulnerability data was retrieved from Yang et al. (2020), based on 104 milk production participants (90 farmers and 14 milk processors) in China. [b] Higher rank of the opportunities and motivations factors indicate higher vulnerability Yang et al. (2020). [c] Higher rank of the control factors indicate more adequate controls and, thus, lower vulnerability Yang et al. (2020). [d] The variable coefficients in the regression vector of the principal component regression (PCR) between the mean rank of the scores of the fraud factors and the percentage of the suspected samples for the four geographical areas.

6.3.4.1. Relation between the origin of the suspected milk and the fraud opportunities and motivations

As presented in Table 6.4, the milk chain actors from the east and central-northern areas, where more suspected UHT milk samples were flagged, stated before that they had poorer business relationships within the supply chain than those in other areas of China (fraud factor 4). Since a good relationship between the actors in the milk supply chain can positively affect information-sharing, this can help to keep the supply chain transparent (Fu, Han, & Huo, 2017), which may additionally play a role in reducing the risk of fraud in certain areas. Conversely, the situation may deteriorate. As a consequence, the regression vector of the PCR showed positive coefficients between the business relationship (fraud factor 4) and the percentage of the suspected samples, indicating that a high rank of the factor would contribute to the higher percentages of suspected samples in these areas.

6.3.4.2. Relation between the origin of the suspected milk and the counteracting controls

The results of our study revealed that two managerial control measures, lack of an ethical code of conduct and lack of integrity screening of the employees, and one technical control measure (lack of fraud contingency plans) related to the higher prevalence of suspected samples from the milk production participants in the east and central-northern areas (Table 6.4). Well-established controls may mitigate against food fraud. However, if they are lacking in combinations with increased opportunities and motivations, companies become increasingly vulnerable to fraud (van Ruth, Huisman, & Luning, 2017).

6.4. Conclusions

The study demonstrated the occurrences of suspected adulterated UHT milk samples in various parts of China and their relationships with previously established fraud vulnerability of businesses operating in those areas. Twelve (out of 52) samples in a market survey were suspects with quality or fraud-related issues, of which one is highly suspected of being adulterated by a dilution with water. The relative prevalence of suspect samples was higher in milks produced in the central-northern and eastern areas than in those produced in the north-western and north-eastern areas, while those of the southern area were in between. The underlying factors contributing to this higher vulnerability are poorer business relationships and a lack of adequate managerial controls.

Acknowledgments

6

Conflicts of interest

The authors declare no conflict of interest.

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.

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

Bergana, M. M., Adams, K. M., Harnly, J., Moore, J. C., & Xie, Z. (2019). Non-targeted detection of milk powder adulteration by 1H NMR spectroscopy and conformity index analysis. *Journal of Food Composition and Analysis, 78*, 49-58.

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.

Chung, I.-M., Park, I., Yoon, J.-Y., Yang, Y.-S., & Kim, S.-H. (2014). Determination of organic milk authenticity using carbon and nitrogen natural isotopes. *Food Chemistry*, *160*, 214-218.

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.

De Longhi, R., Spinardi, N., Nishimura, M. T., Miyabe, M. Y., Aragon-Alegro, L. C., De Rezende Costa, M., & De Santana, E. H. W. (2012). A survey of the physicochemical and microbiological quality of ultra-heat-treated whole milk in Brazil during their shelf life. *International Journal of Dairy Technology*, *65*(1), 45-50.

Du, B., Wen, F., Zhang, Y., Zheng, N., Li, S., Li, F., & Wang, J. (2019). Presence of tetracyclines, quinolones, lincomycin and streptomycin in milk. *Food Control*, *100*, 171-175.

Fox, P. F., McSweeney, P. L., & Paul, L. (1998). *Dairy Chemistry and Biochemistry*. London: Blackie Academic & Professional.

Fu, S., Han, Z., & Huo, B. (2017). Relational enablers of information sharing: Evidence from Chinese food supply chains. *Industrial Management & Data Systems, 117*(5), 838-852.

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*. Waltham, MA: Academic Press.

Guo, J., Liu, X., Xu, A., & Xia, Z. (2010). Relationship of somatic cell count with milk yield and composition in Chinese Holstein population. *Agricultural Sciences in China*, *9*(10), 1492-1496.

6

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.

Handford, C. E., Campbell, K., & Elliott, C. T. (2016). Impacts of milk fraud on food safety and nutrition with special emphasis on developing countries. *Comprehensive Reviews in Food Science and Food Safety*, *15*(1), 130-142.

Hansen, P. W., & Holroyd, S. E. (2019). Development and application of Fourier transform infrared spectroscopy for detection of milk adulteration in practice. *International Journal of Dairy Technology, 72*(3), 321-331.

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.

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.

Kamal, M., & Karoui, R. (2015). Analytical methods coupled with chemometric tools for determining the authenticity and detecting the adulteration of dairy products: A review. *Trends in Food Science & Technology*, *46*(1), 27-48.

Lam, H., Remais, J., Fung, M., Xu, L., & Sun, S.(2013). Food supply and food safety issues in China. *The Lancet*, *381*(9882), 2044-2053.

Li, S. (2016). White paper on China dairy 2016. Retrieved from http://www.sdddc.org/en/download/detail-251.aspx.

Li, S., Min, L., Wang, P., Zhang, Y., Zheng, N., & Wang, J. (2017). Occurrence of aflatoxin M1 in pasteurized and UHT milks in China in 2014-2015. *Food Control*, *78*, 94-99.

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.

Lima, L. S., Rossini, E. L., Pezza, L., & Pezza, H. R. (2020). Bioactive paper platform for detection of hydrogen peroxide in milk. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *227*, 117774.

Liu, N., Koot, A., Hettinga, K., de Jong, J., & van Ruth, S. M. (2018). Portraying and tracing the impact of different production systems on the volatile organic compound composition of milk by PTR-(Quad)MS and PTR-(ToF)MS. *Food Chemistry*, *239,* 201-207.

Ministry of Agriculture P. R. China. (2016). *China Dairy Yearbook*. Beijing, China: China Agriculture Press.

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.

National Standard of the P. R. China. (2010). National food safety standard sterilized milk. Retrieved from http://tradechina.dairyaustralia.com.au/wp-content/uploads/2018/08/GB-25190-2010-National-Food-Safety-Standard-Sterilized-Milk-f1-.pdf.

Nurseitova, M. A., Amutova, F. B., Zhakupbekova, A. A., Omarova, A. S., Kondybayev, A. B., Bayandy, G. A., & Konuspayeva, G. S. (2019). Comparative study of fatty acid and sterol profiles for the investigation of potential milk fat adulteration. *Journal of Dairy Science*, *102*(9), 7723-7733.

Oliveri, P. (2017). Class-modelling in food analytical chemistry: Development, sampling, optimisation and validation issues – a tutorial. *Analytica Chimica Acta*, *982*, 9-19.

PwC. (2018). The ongoing modernisation of China's dairy sector. Retrieved from https://www.pwccn.com/en/food-supply/publications/modernization-of-china-dairy-industry.pdf.

Sánchez, A., Sierra, D., Luengo, C., Corrales, J. C., de la Fe, C., Morales, C. T., & Gonzalo, C. (2007). Evaluation of the MilkoScan FT 6000 milk analyzer for determining the freezing point of goat's milk under different analytical conditions. *Journal of Dairy Science*, *90*(7), 3153-3161.

Schliep, K., Hechenbichler, K., & Lizee, A. (2016). *Kknn: Weighted k-nearest Neighbor*s. R package version, 1. Retrieved from https://cran.r-project.org/web/packages/kknn/kknn.pdf.

Silva, A. F. S., & Rocha, F. R. P. (2020). A novel approach to detect milk adulteration based on the determination of protein content by smartphone-based digital image colorimetry. *Food Control*, *115*, 107299.

Tittlemier, S. A. (2010). Methods for the analysis of melamine and related compounds in foods: A review. *Food Additives & Contaminants: Part A, 27*(2), 129-145.

van Ruth, S. M., Huisman, W., & Luning, P. A. (2017). Food fraud vulnerability and its key factors. *Trends in Food Science & Technology, 67,* 70-75.

Xiong, J., Peng, L., Zhou, H., Lin, B., Yan, P., Wu, W., & Qiu, Y. (2020). Prevalence of aflatoxin M1 in raw milk and three types of liquid milk products in central-south China. *Food Control*, *108,* 106840.

Xiu, C., & Klein, K. K. (2010). Melamine in milk products in China: Examining the factors that led to deliberate use of the contaminant. *Food Policy*, *35*(5), 463-470.

Yang, L., Yang, Q., Yi, M., Pang, Z. H., & Xiong, B. H. (2013). Effects of seasonal change and parity on raw milk composition and related indices in Chinese Holstein cows in northern China. *Journal of Dairy Science*, *96*(11), 6863-6869.

6

Yang, Y., Hettinga, K. A., Erasmus, S. W., Pustjens A. M. & van Ruth, S. M. (2020). Opportunities for fraudsters: When would profitable milk adulterations go unnoticed by common, standardized FTIR measurements? *Food Research International*, *136,* 109543.

Yang, Y., Huisman, W., Hettinga, K. A., Liu, N., Heck, J., Schrijver, G. H., Gaiardoni, L., & van Ruth, S. M. (2019). Fraud vulnerability in the Dutch milk supply chain: Assessments of farmers, processors and retailers. *Food Control*, *95*, 308-317.

Yang, Y., Huisman, W., Hettinga, K. A., Zhang, L., & van Ruth, S. M. (2020). The Chinese milk supply chain: A fraud perspective. *Food Control,113,* 107211.

Yin, S., Li, Y., Xu, Y., Chen, M., & Wang, Y. (2017). Consumer preference and willingness to pay for the traceability information attribute of infant milk formula. *British Food Journal, 119*(6), 1276-1288.

Zhang, W., & Xue, J. (2016). Economically motivated food fraud and adulteration in China: An analysis based on 1553 media reports. *Food Control, 67*, 192-198.

Zhao, Z., Yue, T., Wang, Y., & Peng, B. (2005). Choice of CCP in production process of natural full-cream UHT milk. *China Dairy Industry*, *33*(10), 51.

Zheng, N., Sun, P., Wang, J. Q., Zhen, Y. P., Han, R. W., & Xu, X. M. (2013). Occurrence of aflatoxin M1 in UHT milk and pasteurized milk in China market. *Food Control*, *29*(1), 198-201.

Supplementary material

Table S6.1. The calculations and four levels of adulterants added into the three milk pools (per 100 g) [a].

Adulterant category	Calculation	Adulterant	Level 1 (g)	Level 2 (g)	Level 3 (g)	Level 4 (g)
Protein-rich adulterants	$Weight_{adulterant}$ $= 100\ g \times \left(\dfrac{Protein_{control} \times a\%}{Protein_{adulterant}} \right)$	WMP	1.36	2.72	4.09	5.45
		SMP	0.99	1.98	2.97	3.95
		WPI	0.38	0.75	1.13	1.51
		SOY	0.39	0.78	1.17	1.56
		PEA	0.43	0.85	1.28	1.71
Nitrogen-based adulterants	$Weight_{adulterant}$ $= 100\ g \times \left(\dfrac{Protein_{control} \times a\%}{f \times N_{adulterant}} \right)$	URE	0.12	0.24	0.35	0.47
		MLM	0.08	0.16	0.25	0.33
		AC	0.21	0.42	0.63	0.84
		AS	0.26	0.52	0.78	1.04
		DIC	0.08	0.16	0.25	0.33
Carbohydrate-based adulterants	$Weight_{adulterant}$ $= 100\ g \times \left(\dfrac{TS_{control} \times a\%}{TS_{adulterant}} \right)$	SU	1.30	2.60	3.90	5.20
		GLU	1.30	2.60	3.90	5.20
		FRU	1.30	2.60	3.90	5.20
		LAC	1.30	2.60	3.90	5.20
		MD	1.30	2.60	3.90	5.20
		STA	1.30	2.60	3.90	5.20
		AR	1.53	3.06	4.59	6.12
Preservatives	Not available	CIT	0.05	0.10	0.15	0.20
		CAR	0.05	0.10	0.15	0.20
		BIC	0.05	0.10	0.15	0.20
		FMD	0.05	0.10	0.15	0.20
		PX	0.05	0.10	0.15	0.20
		HYD	0.05	0.10	0.15	0.20
Water	Not available	Water	10.00	20.00	30.00	40.00

[a] $a\%$ stands for the four adulteration levels for the protein-rich, nitrogen-based and carbohydrate-based adulterants, it is equal to 10%, 20%, 30% and 40% for the four levels, respectively. $Protein_{control}$ is the protein content of the control milk samples (3.5% w/w). $Protein_{adulterant}$ is the protein content of the protein-rich adulterant. $N_{adulterant}$ is the nitrogen content of the nitrogen-based adulterant. f is the conversion factor of nitrogen to protein (6.38). $TS_{control}$ is the total solids content of the control milk samples (13.0% w/w). $TS_{adulterant}$ is the total solids content of the carbohydrate-based adulterant. 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.

6

Table S6.2. The amount of adulterants used for the combined adulterations for the three milk pools [a].

Adulterant category	Calculation	Adulterant	Adulterant (g)	Pooled milk (g)	Water (g)
Protein-rich adulterants	$Weight_{adulterant} = \dfrac{\{Protein_{control} \times (1 + a\%) \times (100 + 40)\ g\} - (Protein_{control} \times 100\ g)}{Protein_{adulterant}}$	WMP	13.07	100.00	40.00
		SMP	9.49	100.00	40.00
		WPI	3.61	100.00	40.00
		SOY	3.73	100.00	40.00
		PEA	4.10	100.00	40.00
Nitrogen-based adulterants	$Weight_{adulterant} = \dfrac{\{Protein_{control} \times (1 + a\%) \times (100 + 40)\ g\} - (Protein_{control} \times 100\ g)}{f \times N_{adulterant}}$	URE	1.13	100.00	40.00
		MLM	0.79	100.00	40.00
		AC	2.01	100.00	40.00
		AS	2.49	100.00	40.00
		DIC	0.79	100.00	40.00
Carbohydrate-based adulterants	$Weight_{adulterant} = \dfrac{\{TS_{control} \times (1 + a\%) \times (100 + 40)\ g\} - (TS_{control} \times 100\ g)}{TS_{adulterant}}$	SU	12.48	100.00	40.00
		GLU	12.48	100.00	40.00
		FRU	12.48	100.00	40.00
		LAC	12.48	100.00	40.00
		MD	12.48	100.00	40.00
		STA	12.48	100.00	40.00
		AR	14.68	100.00	40.00

[a] $a\%$ is equal to 40% for the combined adulteration. $Protein_{control}$ is the protein content of the control milk samples (3.5% w/w). $Protein_{adulterant}$ is the protein content of the protein-rich adulterant. $N_{adulterant}$ is the nitrogen content of the nitrogen-based adulterant. f is the conversion factor of nitrogen to protein (6.38). $TS_{control}$ is the total solids content of the control milk samples (13.0% w/w). $TS_{adulterant}$ is the total solids content of the carbohydrate-based adulterant. AC: ammonium chloride; AR: arrowroot powder; AS: ammonium sulphate; DIC: dicyandiamide; FRU: fructose; GLU: glucose; LAC: lactose; MD: maltodextrin; MLM: melamine; PEA: pea protein isolate; SMP: skimmed milk powder; SOY: soy protein isolate; STA: starch; SU: sucrose; URE: urea; WMP: whole milk powder; WPI: whey protein isolate.

**Table S6.3.** The result of the milk compositional features of the control samples for the measured dataset [a].

Sample ID	Protein (% w/w)	Fat (% w/w)	TS (% w/w)	SNF (% w/w)	Lactose (% w/w)	Density (g/L)	FPD (°C)
1	3.61	3.72	13.36	9.71	5.36	1035	0.562
2	3.71	4.43	14.03	9.67	5.21	1033	0.562
3	3.72	4.21	13.98	9.86	5.40	1035	0.570
4	3.74	4.35	13.95	9.66	5.18	1033	0.569
5	3.49	3.88	13.02	9.17	4.96	1032	0.530
6	3.42	3.93	12.86	8.94	4.80	1031	0.516
7	3.55	3.99	13.10	9.13	4.84	1031	0.522
8	3.34	3.60	12.56	8.98	4.91	1031	0.521
9	3.72	4.04	13.75	9.78	5.31	1034	0.576
10	3.33	3.75	12.71	8.99	4.93	1032	0.521
11	3.42	3.97	12.98	9.03	4.88	1031	0.529
12	3.42	3.73	12.82	9.11	4.96	1032	0.538
Pool A	3.69	4.05	13.72	9.73	5.30	1034	0.567
Pool B	3.44	3.75	12.79	9.05	4.88	1031	0.524
Pool C	3.47	3.83	13.04	9.24	5.04	1032	0.544

[a] ID: Identity; FPD: freezing point depression; TS: total solids; SNF: solids non-fat.

6

Table S6.4. The result of the milk compositional features of the control samples for the variance-adjusted dataset [a].

Sample ID	Protein (% w /w)	Fat (% w/w)	TS (% w/w)	SNF (% w/w)	Lactose (% w/w)	Density (g/L)	FPD (°C)
1	3.67	3.48	13.47	10.09	5.66	1038	0.581
2	3.88	4.91	14.81	10.00	5.35	1034	0.581
3	3.89	4.46	14.72	10.39	5.73	1038	0.597
4	3.93	4.74	14.66	9.99	5.29	1034	0.595
5	3.43	3.81	12.80	9.01	4.85	1032	0.517
6	3.29	3.91	12.48	8.55	4.54	1030	0.489
7	3.56	4.02	12.95	8.92	4.62	1030	0.500
8	3.14	3.24	11.88	8.62	4.76	1030	0.499
9	3.90	4.12	14.26	10.23	5.56	1036	0.609
10	3.12	3.54	12.18	8.64	4.79	1031	0.499
11	3.30	3.98	12.72	8.72	4.70	1030	0.515
12	3.30	3.51	12.39	8.88	4.85	1032	0.533
Pool A	3.83	4.14	14.19	10.13	5.54	1036	0.591
Pool B	3.34	3.55	12.34	8.77	4.70	1030	0.505
Pool C	3.39	3.71	12.83	9.14	5.01	1032	0.545

[a] ID: Identity; FPD: freezing point depression; TS: total solids; SNF: solids non-fat.

6
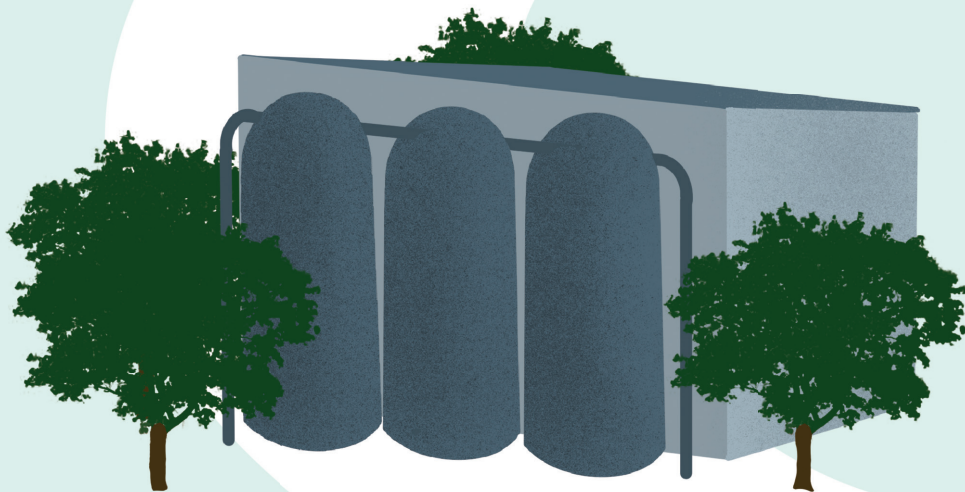
CHAPTER 7

General discussion

7.1. Introduction

Milk supply chains have become more complex over the years because of globalization and industrialization. The products' nature as well as economic and behavioural drivers in combination with this complexity of the supply chain networks make that milk supply chains are susceptible to fraud. Recent milk fraud incidents indicate the continuous threat of food fraud reaching every dining table in the world and highlight the importance to combat these illicit activities. Therefore, this thesis aims at better comprehension of the factors involved in food fraud in liquid milk supply chains. To achieve this objective, Chapters 2 and 3 examined fraud vulnerability of milk supply chains in the Netherlands and China. Next to that, in Chapters 4 and 5 detection approaches for milk adulteration were developed and evaluated. In Chapter 6 the prevalence of milk adulteration in the Chinese market was investigated, applying the detection approaches that were described in Chapter 4. In the current final chapter, the findings from Chapters 2-6 are integrated, to give a comprehensive overview of fraud vulnerability in the Dutch and Chinese milk supply chains, newly developed authentication approaches, and the relationship between fraud vulnerability and actual fraud prevalence in China. Based on these results, fraud mitigation approaches are suggested and recommendations for future research are provided.

7.2. Fraud vulnerability in the Dutch and Chinese milk supply chains

7.2.1. General differences

Individual fraud vulnerability assessments for actors in the Dutch and Chinese milk supply chains were presented in Chapters 2 and 3, respectively. In the current chapter, we compare the vulnerabilities of the two supply chains. Therefore, the results of all Dutch and Chinese food fraud vulnerability assessments were subjected to multiple correspondence analysis (MCA; Fig. 7.1).

First of all, it is striking that the mean scores for the Dutch and Chinese actors do not differ much: The scores (coordinates) for the Dutch actors are 0.09 (first dimension: horizontal axis) and 0.16 (second dimension: vertical axis), and for the Chinese actors -0.03 (first dimension) and -0.06 (second dimension), respectively (means of all scores shown in Fig. 7.1a). However, it is also obvious that the scores of the Chinese actors show a wider spread and are thus varying more than those of the Dutch group. Furthermore, the Dutch and Chinese processors are generally positioned on the left side from the farmers which correlates with a lower vulnerability, but this is more explicit for the Chinese processors than for their Dutch counterparts.
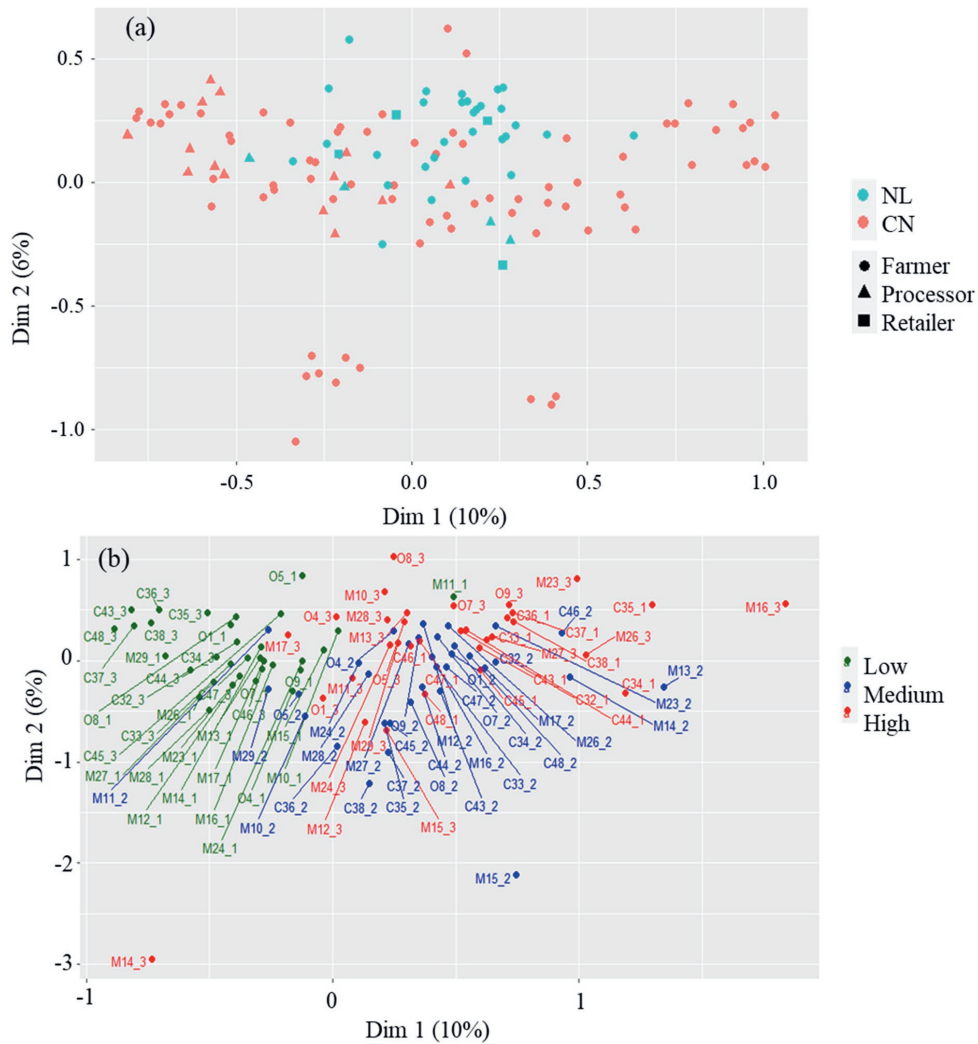
Fig. 7.1. The first two dimensions of multiple correspondence analysis (MCA) on the common fraud factors of food fraud vulnerability assessments: (a) scores plot for farmers (circle), processors (triangle) and retailers (square), in the Netherlands (blue) and China (red); and (b) loadings plot of which high, medium, and low vulnerability scores are coloured red, blue, and green, respectively. Fraud factors not applicable to all tier groups were omitted for this analysis. The fraud factors and their coding are listed in Table 7.1.

Table 7.1. The common fraud factors in the vulnerability assessment for the farmers, processors and retailers, and their coding.

Key elements	Fraud factors
Opportunities	(O1) Ease of adulteration; (O4) Availability of technology and knowledge for adulteration of the final milk product; (O5) Detectability of fraud in the final milk product; (O6) Accessibility to production activities; (O7) Transparency of the chain network; (O8) Relationship within the supply chain; (O9) Historical evidence of milk fraud
Motivations	(M10) Supply and price of milk; (M11) Valuable components or attributes; (M12) Economic health of the own company; (M13) Business strategy of the own company; (M14) Ethical business culture of the own company; (M15) Previous irregularities of the own company; (M16) Corruption level of the country in which the own company is active; (M23) Corruption level of the country in which the customer is active; (M24) Economic health sector; (M26) Sector ethical business culture; (M27) Historical evidence of milk fraud, within sector; (M28) Level of competition in sector; (M29) Price asymmetries
Controls	(C32) Specificity and accuracy of the fraud monitoring system in place for the milk product in the own company; (C33) Verification of fraud monitoring system for the milk product in the own company; (C34) Accuracy of the information system wrt mass balance control in the own company; (C35) Extensiveness of the tracing and tracking system in the own company; (C36) Application of integrity screening of employees in the own company; (C37) Strictness of the ethical code of conduct in the own company; (C38) Support of a whistle blowing system in the own company; (C43) Social control and transparency across the chain network; (C44) Established guidance for fraud prevention and control in the sector; (C45) Specificity of the national food policy; (C46) Strictness of law enforcement in the local chain; (C47) Strictness of law enforcement in the international chain; (C48) Availability of a fraud contingency plan

7.2.2. Differences between dairy farms

The mean scores for the Dutch and Chinese dairy farms are fairly similar: The scores (coordinates) for the Dutch farmers are 0.11 and 0.21, and for the Chinese farmers 0.03 and -0.08, respectively (means of all scores shown in Fig. 7.1a for first and second dimension, respectively). However, the standard deviations (SD) of the scores of the Chinese farmers along the first and second dimension in Fig. 7.1a (0.51 and 0.40, respectively) are much larger than those of their Dutch counterparts (0.20 and 0.17, respectively), indicating that scores vary more among the Chinese farmers. This result implies more extreme fraud vulnerability profiles of Chinese actors, both when it comes to the lower and higher end of the scale. This difference in variation between the farmers of the two countries may be due to several reasons. On the one hand, the Netherlands is a relatively small-sized country with little geographical and cultural variation. Together with the Dutch dairy chain, dairy farms have gradually developed over many decades, with many self-imposed regulations, leading to little variation in practices (Bijman, 2018; van Asseldonk & Velthuis, 2014). On the other hand, the Chinese dairy farms are highly scattered over a much larger territory, resulting in more variation in culture, climate, economic environment, crop production, etc. (PwC, 2018; Ma, Oxley, Rae, Fan, Huang, &

Rozelle, 2012). These variations may underlie the larger variability in fraud vulnerability among the Chinese farms.

Table 7.2. Mean ranks of scores of the fraud factors from the food fraud vulnerability assessments showing significant difference* between the dairy farmers from the Netherlands (NL) and China (CN).

Fraud element	Fraud factor and coding	NL ($n = 30$)	CN ($n = 90$)	p value
Opportunities [a]	O4 Availability of technology and knowledge for adulteration of the final milk product	**77** [c]	55	0.001
	O6 Accessibility to production activities	**89**	51	< 0.001
	O8 Relationship within the supply chain	44	**66**	0.001
	O9 Historical evidence of milk fraud	**73**	56	0.003
Motivations [a]	M11 Supply and price of milk	**73**	56	0.015
	M12 Economic health of the own company	47	**65**	0.006
	M16 Corruption level of the country in which the own company is active	42	**67**	< 0.001
	M17 Financial pressure from the supplier	39	**68**	< 0.001
	M18 Customer's economic health	47	**65**	0.002
	M22 Victimization of the customer	**72**	57	0.016
	M23 Corruption level of the country in which the customer is active	**76**	55	0.001
	M28 Level of competition in sector	46	**66**	0.004
	M29 Price asymmetries	51	**64**	0.036
Controls [b]	C35 Extensiveness of the tracing and tracking system in the own company	**79**	54	< 0.001
	C36 Application of integrity screening of employees in the own company	30	**70**	< 0.001
	C37 Strictness of the ethical code of conduct in the own company	25	**72**	< 0.001
	C38 Support of a whistle blowing system in the own company	28	**71**	< 0.001
	C40 Specificity and accuracy of the customer fraud monitoring system	**73**	56	0.007
	C42 Extensiveness of the customer tracking and tracing system	**71**	57	0.039
	C45 Specificity of the national food policy	**77**	55	< 0.001
	C48 Availability of a fraud contingency plan	40	**67**	< 0.001

* Mann-Whitney U test, $p < 0.05$. [a] Higher rank of the opportunities and motivations indicate higher vulnerability. [b] Higher rank of the controls indicate more adequate controls, and thus lower vulnerability. [c] Higher ranks for each factor are presented in bold letter type.

To zoom in on the differences between the farmers of the two countries, the results for the individual factors are statistically compared (Table 7.2). It appears that for the factor regarding opportunities, motivations, and controls there are always some of these factors presenting higher vulnerability for the Dutch farmers and some for the Chinese farmers. For the opportunities four out of the seven factors show significant differences, with three indicating

7

higher vulnerability for the Dutch and one for the Chinese farmers. Similarly, nine out of the 19 motivation factors reveal significant differences, with three showing higher vulnerability levels for the Dutch and six for the Chinese farmers. Finally eight of the 17 control factors demonstrate significant differences, with the highest vulnerability recorded half-half for Dutch and Chinese farmers. These differences appear fairly balanced between the Dutch and Chinese farmers, except for the economic drivers, which show some over-representation in terms of vulnerability in the Chinese farmers' group.

Different scores in opportunities might be caused by the different cultures of the two countries. Most Dutch farms were open to public, because their owners believed the farm to be a bridge between the people and nature. On the contrary, most of the Chinese farmers considered milk production as the only purpose to run a farm, so they limited the access to their farms only to the authorized employees (Li, 2016). Several reasons caused the higher vulnerability levels regarding motivation-related factors for the Chinese farmers. On the one hand, the unbalanced development of dairy production and processing in the Chinese dairy industry may have caused the farmers to become price-takers. The dairy processors determine the farm gate milk price, and own profit maximization drove them to limit raw milk price marginally over the cost of production in the past decade (Li, 2016). As a result, the Chinese farmers are and were under large financial pressure, and the economic condition of farms depends largely on their customers. Moreover, the loss of confidence in local dairy products after a number of milk incidents drove the consumers to purchase imported alternatives (El Benni et al., 2019; Kendall et al., 2019; Li, Sijtsema, Kornelis, Liu, & Li, 2019), which increased the level of competition. On the other hand, the Dutch farmers benefit from their involvement in dairy cooperatives, therefore, the economic drivers were rated at a lower vulnerability level by this group. The results also showed that the Dutch farmers reported less adequate managerial controls, while the Chinese farmers reported less adequate technical controls. As many Dutch farmers are family businesses, requiring no or only a few employees, the internal managerial controls were not extensive within Dutch farms, which is very common in small-sized organisations. Unlike Dutch farms, the Chinese farms were still at the stage of developing and transforming from household husbandry to intensive breeding, and they are still in the process of establishing technical controls (Qian, Song, Hu, & Ying, 2018).

Furthermore, it shows clearly that there are groups of farmers with extremely low or high levels of fraud vulnerability (Fig. 7.1a). A group of very low vulnerability actors can be distinguished at the far left hand side of the plot. This group, consisting of Chinese farmers primarily, has provided low vulnerability ratings for nearly all questions (loadings plot Fig. 7.1b). This group is either very well protected or may have provided the more socially desirable answers. It was shown that, in social surveys, respondents tend to give socially desirable answers instead of choosing responses that reflect their true feelings (Chung & Monroe, 2003). Another interesting group is the one at the far right hand side of the plot, which indicated higher vulnerability levels for most factors. The loadings plot (Fig. 7.1b) reveals that this holds in particular with regard to lack of control measures. It was previously shown that multiple factors, including organization size, perception and attitude towards regulations or standards, reputation,

expected cost and benefit, and the destination market, can all influence the establishment of controls within businesses (Jin & Zhou, 2011). Any of these factors might cause control flaws for this particular group. Finally, there is a third interesting group at the bottom of the plot consisting of ten Chinese farmers. They rated vulnerability related to ethical business culture of the own company (M14) and previous irregularities of the own company (M15) high. This may indicate vulnerabilities from within the company rather than from the outside. A similar group with vulnerabilities from within the own organisation was seen before in organic food supply chains (van Ruth & de Pagter-de Witte, 2020). The particular group located at the bottom in Fig. 7.1a comprises mainly farms from the central north of China, which is the area where the melamine incidents started and became prevalent (Wu, Zhao, & Li, 2009). This might explain that the previous irregularities (M15) were rated at a higher vulnerability level by this particular group of farmers.

7.2.3. Differences between processors

As shown in Fig. 7.1a, the Chinese processors are positioned somewhat towards the left compared to the Dutch processors. These Chinese processors reported lower fraud vulnerability levels on average (mean score -0.41 along first dimension) compared with their Dutch counterparts (mean score -0.04 along first dimension). In contrast to the farmers, the difference in variation between the processors of the two countries are more similar (SD for the Dutch and Chinese processors are 0.35 and 0.26, respectively, along the first dimension, and are 0.15 and 0.18, respectively, along the second dimension).

It is remarkable that the Dutch and Chinese processors did not show any significant differences in their vulnerability responses regarding opportunities-related and motivations-related factors (Table 7.3) except for the historical evidence of milk fraud (O9). The differences between the Dutch and Chinese processors relate predominantly to differences in adequacy of the control measures: Seven out of the 19 control factors present significant differences in vulnerability level (Table 7.3). Both the Dutch and Chinese processors have reported fairly adequate controls. However, the (even) more extensive Chinese control schemes reported by the Chinese processors may have been triggered by the melamine scandal, which led to a comprehensive reform of the Chinese food safety regime (Pei et al., 2011). The Chinese government issued more than 20 laws and regulations to strengthen milk control (Wu et al., 2018), for instance the Administration for Quality Supervision, Inspection and Quarantine (AQSIQ) issued the Examination Rules for Production Licensing Conditions of Dairy in 2010 (AQSIQ, 2010). These regulations highlight the quality controls required for the processors, for instance with regard to establishing track and trace systems. After issuing new regulations, regulators began to re-check the application of production licensing. After this re-evaluation, 688 out of 1176 dairy enterprises passed (Wu et al., 2018). This may be the reason why Chinese processors reported to have implemented very adequate controls.

7

Table 7.3. Mean ranks of scores of the fraud factors from the food fraud vulnerability assessments showing significant difference* between the dairy processors from the Netherlands (NL) and China (CN).

Fraud element	Fraud factor and coding	NL ($n = 4$)	CN ($n = 14$)	p value
Opportunities [a]	O9 Historical evidence of milk fraud	**15 [c]**	8	0.001
Controls [b]	C30 Specificity and accuracy of the fraud monitoring system in place for incoming milk in the own company	3	**11**	0.001
	C31 Systematics and autonomy of verification of the fraud monitoring system for incoming milk in the own company	4	**11**	0.015
	C32 Specificity and accuracy of the fraud monitoring system in place for the milk product in the own company	3	**11**	0.002
	C33 Systematics and autonomy of verification of fraud monitoring system final product	5	**11**	0.028
	C44 Established guidance for fraud prevention and control in the sector	4	**11**	0.014
	C46 Strictness of law enforcement in the local chain;	4	**11**	0.004
	C48 Availability of a fraud contingency plan	3	**11**	0.001

* Mann-Whitney U-test, $p < 0.05$. [a] Higher rank of the opportunities indicate higher vulnerability. [b] Higher rank of the controls indicate more adequate controls, and thus lower vulnerability. [c] Higher ranks for each factor are presented in bold letter type.

7.3. Approaches for detection of milk adulteration

Fraud detection by analytical means is an important control measure and at the same time it helps to limit the impact of food fraud. Therefore, in this thesis, Chapters 4-6 focused on milk fraud authentication techniques. An attempt was made to develop univariate and multivariate approaches to detect a wide variety of adulterants using the data of milk quality parameters, that are gathered usually by standardised Fourier transform infrared spectroscopy (FTIR) analysis for payment schemes. Furthermore, novel sound-based (ultrasonic) measurements were explored as a means for rapid and non-destructive characterisation of milks and creams, and detection of adulteration (Chapter 5).

7.3.1. Models based on data from routine compositional analysis

7.3.1.1. Natural variation of commercial Dutch and Chinese milk samples

Understanding the natural variation in composition of non-adulterated milk samples is the first step for its authentication. These were measured for both Dutch and Chinese ultra-high temperature (UHT) milk samples, and slight differences in composition between the Dutch and Chinese UHT milk products were discovered. The fat content and lactose content of the Dutch UHT milk (3.8% w/w and 4.7% w/w, respectively) were significantly lower (t-test, $p < 0.05$)

than those of the Chinese ones (4.0% w/w and 5.0% w/w, respectively), whereas the protein contents of the milk from the two countries did not differ significantly (Chapters 4 and 6). It is also noted that the natural variation is larger among the Chinese samples (coefficient of variances (CVs) of protein, fat, and lactose contents were 4.2% w/w, 6.1% w/w, and 4.1% w/w, respectively) than that among the Dutch samples (CVs of protein, fat, and lactose contents were 4.1% w/w, 2.8% w/w, and 1.3% w/w, respectively). The result of milk composition is generally in line with the previous studies (Guo, Liu, Xu, & Xia, 2010; Heck, van Valenberg, Dijkstra, & van Hooijdonk, 2009; Walstra, Wouters, & Geurts, 2005). The raw milk in China may vary more due to the large-sized area with many climatological, economic, cultural, etc. differences compared to the Netherlands. The differences in milk composition and variation may also be caused by different treatments during processing. In the Netherlands, only the fat content is standardised during UHT milk processing. In China, however, flash evaporation is sometimes applied in processing of premium UHT milk to achieve higher protein, fat, and lactose contents (Zhao, Yue, Wang, & Peng, 2005). This processing step could, hence, result in larger variation in composition of Chinese milk products.

7.3.1.2. Impact of common adulterants on apparent milk composition

The impact of common adulterants on apparent milk composition for the Dutch commercial milks was presented in Chapter 4, and the corresponding impact for the Chinese commercial milks were determined based on the results in Chapter 6. Although the fat and lactose contents differed significantly between the Dutch and Chinese commercial UHT milk (see 7.3.1.1), the impact of common milk adulterations on the Dutch and Chinese milk was consistent. The slopes of the linear regression models, which were performed based on the apparent readings and adulteration levels for the common adulterants, presented no significant differences between the Dutch and Chinese milk samples. These results show the robustness of the chosen approach. Despite the different milk composition of the control groups, the routine compositional analysis successfully allowed the detection of common adulterants. It is, therefore, promising to be applied to, for instance milk from different countries, although separate authentication models may need to be developed.

7.3.1.3. Differences between data analysis approaches

Different approaches were used for distinction of genuine and artificially adulterated milk, and their selectivity compared. On the one hand, the measured dataset was used as baseline and on the other hand the same dataset but with its variance adjusted to values common in practice. Furthermore, a univariate approach was applied, which considered the individual, measured characteristics, but also a multivariate approach was considered. Below these approaches are compared for the Dutch and Chinese milk sample sets.

Measured dataset versus variance-adjusted dataset. The result of the measured dataset strictly represented the variation of the selected genuine (unadulterated) samples. The findings in Chapters 4 and 6 show that the boundaries determined by the measured dataset identified

7

more abnormal samples, compared to the variance-adjusted dataset for both the Dutch and Chinese sample set. However, the limited number of samples, covering only a limited number of locations with a single production season, obviously does not cover variation encountered in practice. Therefore, the risk of flagging genuine samples as suspect samples is high using this approach. Contrary to the measured dataset, the variance-adjusted dataset represented a larger variation which appeared to be closer to the variation that occurs in the real world. Consequently, the risk of misclassification of genuine samples was lower using this dataset. At the same time, however, these looser boundaries decreased the specificity for identification of adulterated samples too. This all holds for both the Dutch and Chinese milk sample sets. The selection of appropriate thresholds will, therefore, inevitably always be a trade-off (van Cauter, 1988), as is true for any screening test, and should depend on the users' requirements in practice.

Univariate versus multivariate approaches. It was surprising that for the datasets of both the Dutch and Chinese UHT milk samples, the univariate approaches flagged adulterations of milk samples at lower concentration levels than the multivariate approaches (Chapters 4 and 6). The univariate approach also indicates directly which component(s) is/are outside the boundaries, thus making it easier to identify the specific reason for the potential food fraud or other quality issues. However, it is also easier for fraudsters to circumvent the test and to ensure that values are met artificially, since they know which parameters are tested and which limits are used. On the contrary, it is much harder to adulterate a product in a way that its multivariate pattern is accepted by a model based on genuine products (Leardi, 2018). One-class models, as used in this thesis, have the advantage of detection of a wide range of adulterations at once. Even those, that may be invented by fraudsters in the future. The trade-off is that it may be somewhat less selective for a particular adulterant as shown in this study. Choices need to be made depending on the specific aims to conduct tests. If speed is of key priority, a broad anomaly test may be the preferred way. If one is interested more in particular adulterations, it is better to choose the higher selectivity for that specific adulteration and use a univariate approach. The great advantage of all is that it does not affect the actual analytical measurement, it is still based on the FTIR quality testing that is routinely conducted for payment schemes. It is just a matter of selection of calculations or even concurrent calculations. It is important to keep in mind that in the current thesis a limited number of classifiers (algorithms) have been applied in the multivariate analyses. Many additional algorithms are available for developing classification model, for instance support vector machines (SVM), artificial neural networks (ANN), partial least squares discriminant analysis (PLS-DA), etc. With enlargement of the database in the future, other algorithms may be tested as well to improve performance of the test.

When comparing the datasets of the Dutch and Chinese milk samples, all the approaches mentioned above performed better in flagging adulterated samples with the dataset of the Dutch milk samples. This can be explained by the fact that the variation in different compositional parameters was smaller for the Dutch than Chinese milk samples, for instance, the CVs of fat and lactose contents of Dutch milks were less than half of those of the Chinese milks (see section 7.3.1). Consequently, the boundaries determined for the Dutch milk samples were

stricter, which resulted in a higher number of adulterated milk samples to be identified. It is completely understandable that the milks from a large sized country as China, which varies enormously geographically, also vary more in composition than those of a small-sized country with many standardised procedures, like the Netherlands. Furthermore, the difference in milk composition between premium and normal milk also increase the variation for the Chinese milks. In practice, the selectivity of the models for Chinese milk may be improved by development of mathematical models for milks from particular geographical areas in China.

7.3.2. Novel methodology: ultrasonic analysis

Ultrasound was explored for its ability to differentiate milk and cream samples with various fat contents (Chapter 5). The developed ultrasonic echo-pulse system presented a good performance in distinguishing samples with different fat levels. In addition, as a rapid and non-destructive method, the ultrasonic technique showed some potential in detecting milk dilution. However, compared with the routine compositional analysis approach presented in section 7.3.1, detection of dilution was worse. Although the velocity can be related to some compositional parameters of milk, milk dilution resulted in multiple impacts on velocity. These multiple impacts made it difficult to detect lower dilution levels. However, one has to keep in mind that this has been only the first attempt to apply ultrasonic analysis for milk authentication, and the technology development is at a very different stage than the routine compositional FTIR analysis. Sound-based analysis has its advantages and potential in regard to rapid, non-destructive analysis of food products, but this will require further steps towards development, optimization, and implementation.

7.4. Milk fraud prevalence in Chinese market and its relation with fraud vulnerabilities

Milk fraud prevalence was examined in the Chinese market by collection of 52 UHT milks which were tested applying the approaches based on routine milk composition parameters determined by routine compositional analysis as described in Chapter 4. Among the 52 UHT milk products, twelve suspicious samples were detected. The protein contents of all twelve suspect samples were below the thresholds determined, the fat contents of six of them were below the fat content threshold, and for five lactose contents below the thresholds were measured as well. Among the twelve suspect samples, four were produced in the central-northern area of China, three in the eastern area, three in the southern area, and two in the north-western area (Chapter 6). Proportionally, more suspected samples were discovered among the products produced in the central-northern and eastern areas. This is consistent with the results in Fig. 7.1a, which showed that these two groups were positioned in the higher vulnerability areas (top right and bottom). This prevalence is also in line with the findings of previous investigations, that in the eastern and central-northern regions in China, higher numbers of food fraud incidents have been reported, compared with those in the north-western and southern areas (Liu & Ma, 2016; Zhang & Xue, 2016). Some factors appeared particularly relevant, such as poor business relationships and lack of adequate managerial controls (Chapter 6).

7

7.5. Practical implications for fraud mitigation

The findings from the fraud vulnerability assessments in this thesis underpinned factors for the development of fraud mitigation plans for the Chinese milk supply chains, based on which can be suggested that a proper fraud mitigation strategy should concern several aspects. Firstly, improvement of milk authentication techniques is crucial for fraud mitigation. These techniques include not only confirmatory analyses, but also rapid and portable screening analyses that can be positioned at specific vulnerable point along complex supply chain. Secondly, a well-established price monitoring system within the Chinese industry can discover anomalous price fluctuation, which would be helpful to counteract the increased economic drivers. If any unnatural price fluctuations occur, the government may utilize price control schemes for stabilising prices, thereby stimulating food integrity. Thirdly, sufficient managerial controls within the organizations are related to lower milk fraud prevalence, and would thereby play a vital role in fraud mitigation. It is suggested that Chinese dairy farmers should establish more adequate internal controls to counteract the enhanced opportunities and motivations. Finally, food legislation and governmental regulatory systems play an important role in ensuring food integrity, the regulators are, hence, suggested to pay more attention to food fraud prevention and enforce supervision.

Fraud monitoring by analytical means is an important fraud control measure and limits the impact of the actions of fraudsters when fraud is detected at an early stage. This requires proper milk authentication methods. The standardised FTIR instrument is widely used for routine milk quality control and determination of milk composition, the results of which are also used in farmers' payment schemes. Considering that no extra workload for measurement is required if the data would also be used for milk fraud detection, it is suggested to use the data from the routine compositional FTIR measurements for this purpose, as it allows rapid detection of common milk adulterants with little effort.

Comprehension of fraud vulnerability is the first step to combat food fraud. Food fraud vulnerability assessments were shown to be an efficient tool to identify specific vulnerabilities. It was previously shown that a problem in any link of the supply chain can influence the food integrity of the entire chain (Wu et al., 2018). It is, therefore, suggested to assess the fraud vulnerability of all food businesses which are part of a specific supply chain in an integral way. This thesis showed that fraud vulnerability was correlated with the actual fraud prevalence. Consequently, proper counteracting actions should be taken, based on the outcomes of the fraud vulnerability assessments. Businesses of low vulnerability to fraud require relatively simple mitigation plans, whereas for the weaker points in the supply chain, more elaborate mitigation plans would be required to reduce or eliminate risks.

7.6. Final conclusions

This thesis aimed at elucidating and comparing food fraud vulnerabilities in the Dutch and Chinese milk supply chains, their underlying factors, and the relationship with fraud prevalence. According to the results described in the previous sections, the following conclusions can be drawn from this thesis:

In very general terms, the milk supply chains in both the Netherlands and China showed low to medium fraud vulnerability. However, considerable differences between groups existed. The vulnerability of the Dutch processors was rated lower than those of the Chinese processors, whereas the Dutch and Chinese farmers showed similar vulnerability on average. The variation among the processors' groups of both countries was fairly similar. In contrary, although the Dutch farmers showed a reasonably uniform vulnerability profile, Chinese farmers showed a much wider range of vulnerability levels, i.e. with extremes at the lower and higher end of the scale. Those with high vulnerability levels were located in particular areas in China.

The approach focusing on models based on data from routine compositional analysis succeeded in adequate detection of most of the protein-rich, nitrogen-based, and carbohydrate-based milk adulterants at low concentration levels. However, its capability to detect preservatives was relatively low. The developed ultrasonic technique is a novel method for dairy characterisation using sound propagation properties of dairy samples. As a rapid, non-destructive method, ultrasound showed potential for detection of milk dilution, but requires further development.

The food fraud vulnerability profiles revealed weak points in the supply chains, whereas the detection techniques discovered food fraud prevalence. The investigation into the UHT milk products in the Chinese market showed that a higher proportion of suspect milk was found among samples produced in the central-north and eastern areas, compared with those from north-western and north-eastern areas. The higher prevalence of milk fraud in certain areas correlated well with elevated fraud vulnerability.

Finally, this thesis resulted in new knowledge about food fraud vulnerabilities and underlying factors across nodes in two very different milk supply chain networks, as well as their relationship with fraud prevalence, which was determined by newly developed authentication approaches. This all contributes to the comprehension of the ecology of food fraud which is essential to design appropriate food fraud interventions in the future.

7.7. Research limitations and recommendations

Several limitations are considered for this thesis.

Firstly, it is known that a social desirability bias might be present for sensitive topics. Desirability bias increases the preference of the respondents to project a favourable image to others and thereby avoid embarrassment (Fisher, 1993), especially for some sensitive questions.

7

Considering the food fraud vulnerability assessment comprises a number of sensitive questions, the desirability bias may exist for the assessment. For instance, for the questions regarding the ethical business culture and previous irregularities of the own company in the FFVA survey, the respondents might have chosen a more positive answer. Moreover, the social desirability has different impacts on respondents from different cultural backgrounds (Johnson & Van de Vijver, 2003).The difference in social desirability between the Dutch and Chinese respondents is not yet explained clearly by academics, nevertheless, it needs to be considered when comparing the perceived fraud vulnerability in the milk supply chains in the two countries.

Secondly, a lack of accuracy in the translation of a question regarding historical evidence of fraud into Chinese might have caused some Chinese respondents to have interpreted this question incorrectly, as was discussed in Chapter 3. The original question was about historical evidence of general milk fraud, however, the translated question was easy to be interpreted as asking about milk fraud within the own company. This lack of accuracy in translation led to an uncertainty in the result for this specific question.

Thirdly, a selection bias might exist in the recruitment of participants for fraud vulnerability assessment. The results of the vulnerability assessment were obtained from those who agreed to join the survey. The reasons why the dairy chain actors refused to participate were not investigated. It is recommended to evaluate the impact of this selection bias in future.

The last consideration that can be raised is the size of the dataset of authentic milk samples that was used. An extensive dataset, which covers the natural variation occurring in unadulterated milk samples would be helpful to consolidate the classification model. This would be of interest for improvement of the statistical models for milk authentication.

References

Administration for Quality Supervision, Inspection and Quarantine (AQSIQ) of P. R. China (2010). The examination rules for production licensing conditions of dairy. Retrieved from http://www.gov.cn/gzdt/2010-11/05/content_1738847.htm.

Bijman, J. (2018). Exploring the sustainability of the cooperative model in dairy: The case of the Netherlands. *Sustainability*, *10*(7), 2498.

Chung, J., & Monroe, G.S. (2003). Exploring social desirability bias. *Journal of Business Ethics, 44*, 291-302.

El Benni, N., Stolz, H., Home, R., Kendall, H., Kuznesof, S., Clark, B., & Stolze, M. (2019). Product attributes and consumer attitudes affecting the preferences for infant milk formula in China – a latent class approach. *Food Quality and Preference*, *71*, 25-33.

Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research, 20*(2), 303-315.

Fu, S., Han, Z., & Huo, B. (2017). Relational enablers of information sharing: Evidence from Chinese food supply chains. *Industrial Management & Data Systems*, *117*(5), 838-852.

Guo, J., Liu, X., Xu, A., & Xia, Z. (2010). Relationship of somatic cell count with milk yield and composition in Chinese Holstein population. *Agricultural Sciences in China*, *9*(10), 1492-1496.

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.

Huo, B., Zhao, X., & Zhou, H. (2014). The effects of competitive environment on supply chain information sharing and performance: An empirical study in China. *Production and Operations Management*, *23*(4), 552-569.

Jin, S., & Zhou, J. (2011). Adoption of food safety and quality standards by China's agricultural cooperatives. *Food Control, 22*(2), 204-208,

Johnson, T. P., & Van de Vijver, F. J. (2003). Social desirability in cross-cultural research. *Cross-Cultural Survey Methods, 325*, 195-204.

Kendall, H., Kuznesof, S., Dean, M., Chan, M., Clark, B., Home, R., & Frewer, L. (2019). Chinese consumer's attitudes, perceptions and behavioural responses towards food fraud. *Food Control*, *95,* 339-351.

Leardi, R. (2018). Chemometric methods in food authentication. In D. Sun (Ed.), *Modern Techniques for Food Authentication*. Waltham, MA: Academic Press.

Li, S. (2016). White paper on China dairy 2016. Retrieved from http://www.sdddc.org/en/download/detail-251.aspx.

7

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.

Liu, P., & Ma, L. (2016). Food scandals, media exposure, and citizens' safety concerns: A multilevel analysis across Chinese cities. *Food Policy*, *63,* 102-111.

Ma, H., Oxley, L., Rae, A., Fan, C., Huang, J., & Rozelle, S. (2012). The evolution of productivity performance on China's dairy farms in the new millennium. *Journal of Dairy Science, 95*(12), 7074-7085.

Pei, X., Tandon, A., Alldrick, A., Giorgi, L., Huang, W., & Yang, R. (2011). The China melamine milk scandal and its implications for food safety regulation. *Food Policy*, *36*(3), 412-420.

PwC. (2018). The ongoing modernisation of China's dairy sector. Retrieved from https://www.pwccn.com/en/food-supply/publications/modernization-of-china-dairy-industry.pdf.

Qian, Y., Song, K., Hu, T., & Ying, T. (2018). Environmental status of livestock and poultry sectors in China under current transformation stage. *Science of the Total Environment, 622*, 702-709.

Schein, E. H. (2010). *Organizational Culture and Leadership* . San Francisco, CA: Jossey-Bass.

Transparency International (2018). Corruption perceptions index 2018. Retrieved from https://www.transparency.org/en/cpi/2019/results.

van Asseldonk, M. A. P. M., & Velthuis, A. G. J. (2014). Risk-based audit selection of dairy farms. *Journal of Dairy Science, 97*(2), 592-597.

Van Cauter, E. (1988). Estimating false-positive and false-negative errors in analyses of hormonal pulsatility. *American Journal of Physiology-Endocrinology and Metabolism*, *254*(6), E786-E794.

van Ruth, S. M., & de Pagter-de Witte, L. (2020). Integrity of organic foods and their suppliers: Fraud vulnerability across chains. *Foods, 9*(2)*,* 188.

van Ruth, S. M., Huisman, W., & Luning, P. A. (2017). Food fraud vulnerability and its key factors. *Trends in Food Science & Technology, 67*, 70-75.

van Ruth, S. M., Luning, P. A., Silvis, I. C. J., Yang, Y., & Huisman, W. (2018). Differences in fraud vulnerability in various food supply chains and their tiers. *Food Control*, *84*, 375-381.

Walstra, P., Wouters, J. T., & Geurts, T. J. (2005). *Dairy Science and Technology.* Boca Raton, FL: CRC press.

Wu, X., Lu, Y., Xu, H., Lv, M., Hu, D., He, Z., & Feng, Y. (2018). Challenges to improve the safety of dairy products in China. *Trends in Food Science & Technology*, *76*, 6-14.

Wu, Y., Zhao, Y., & Li, J. (2009). A Survey on occurrence of melamine and its analogues in tainted infant formula in China. *Biomedical and Environmental Sciences, 22*(2), 95–99.

Zhang, W., & Xue, J. (2016). Economically motivated food fraud and adulteration in China: An analysis based on 1553 media reports. *Food Control*, *67*, 192-198.

Zhao, Z., Yue, T., Wang, Y., & Peng, B. (2005). Choice of CCP in production process of natural full-cream UHT milk. *China Dairy Industry*, *33*(10), 51.

7
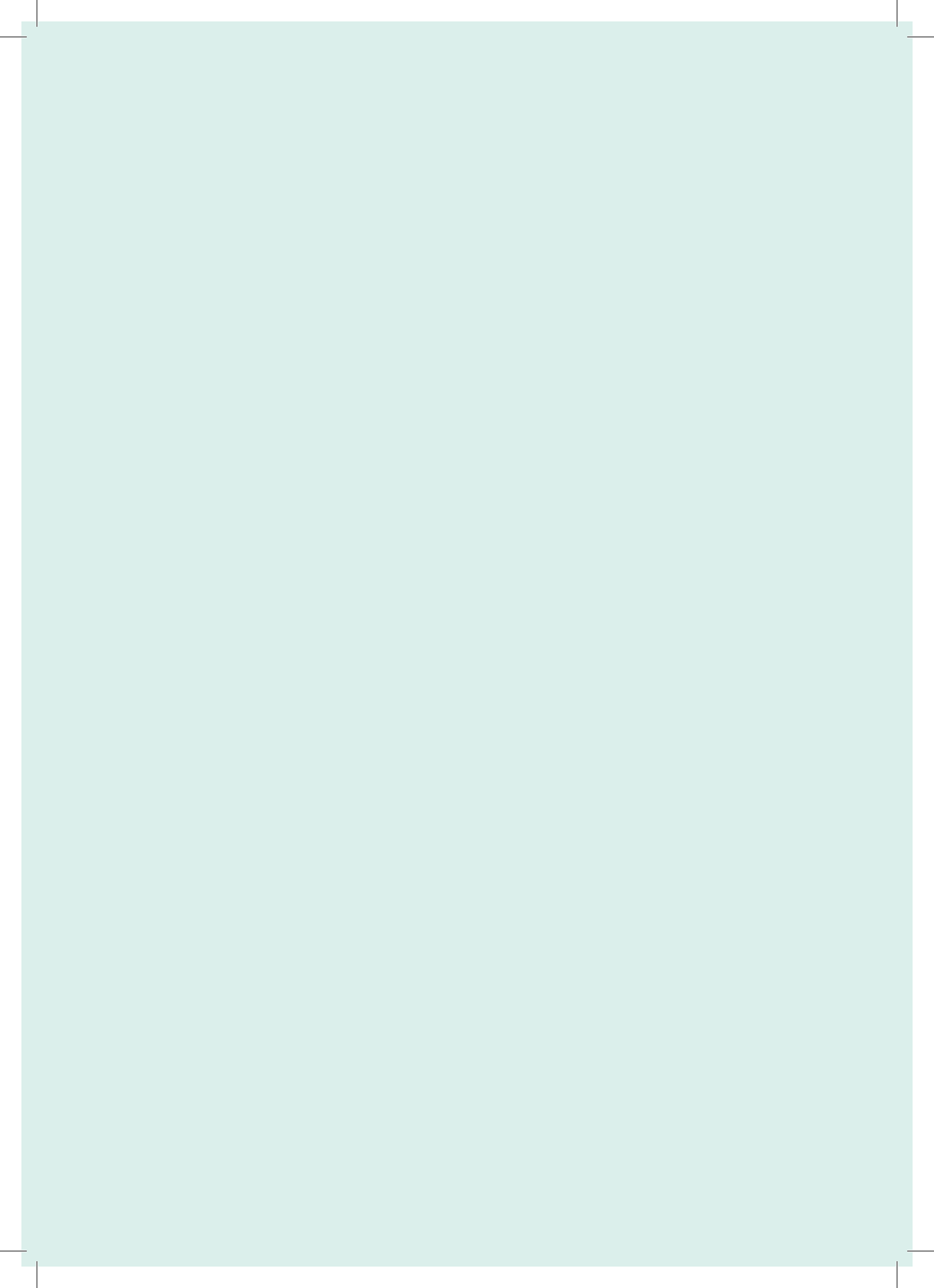
Summary

Summary

Recent food fraud incidents have highlighted the importance of fighting food fraud and thereby ensuring food integrity. Milk is reported as one of the most commonly adulterated foods in the world, hence there is a strong demand for the assurance of its integrity. Both technical and managerial approaches can help to combat food fraud and ensure the integrity of the milk supply chain. The Dutch and Chinese milk supply chains are different with regard to their structure, history of development, and historical evidence of fraud. Hence, this thesis aimed to elucidate and compare fraud vulnerabilities in the Dutch and Chinese milk supply chains, their underlying factors, and the relationship with fraud prevalence.

Fraud vulnerabilities in the Dutch and Chinese supply chains were examined, and the similarities and differences between the tier groups were identified (Chapters 2 and 3). In general, both supply chains were rated as low to medium vulnerable to food fraud by its actors. The variation in fraud vulnerability among different tier groups, i.e. farmers, processors, and retailers, was relatively small in the Netherlands. When comparing Dutch farms of different management types, the organic farms appeared more vulnerable than their non-organic counterparts (Chapter 2). Unlike the Dutch milk supply chain, the difference in fraud vulnerability level between Chinese farmers and processors was larger. The farmers were the weaker link within the Chinese milk supply chain (Chapter 3). On average, Dutch and Chinese farmers reported a similar level of fraud vulnerability, but the Chinese farmers varied much more in their levels. The Dutch processors appeared slightly more vulnerable on average than their Chinese counterparts, but variation in both groups was similar.

Following the profiling of fraud vulnerability of the milk supply chains, analytical methods for detection of milk adulteration were developed. Considering the diversity in milk composition, the authentication approaches for the Dutch and Chinese commercial milks were developed separately, but along the same principles. The latter one was then used for checking the prevalence of suspicion for milk fraud in the Chinese market. The milk compositional parameters measured by routine analysis, which is based on Fourier transform infrared (FTIR) spectroscopy, were used for the development of the milk authentication approaches. Both univariate and multivariate analyses were conducted, and their performance was compared. Using the developed approach, most of the milk samples (more than 90%) that were artificially adulterated with protein-rich, nitrogen-rich, and carbohydrate adulterants could be successfully identified as adulterated. Comparing the univariate and multivariate approaches, the univariate approach performed better in detection of common milk adulteration, missing only one adulterated sample among the protein-rich, nitrogen-rich, and carbohydrate adulterations (Chapter 4). In addition to the routine analysis, a novel, rapid and non-destructive technique, based on an ultrasound echo-pulse system, was explored for the characterisation of dairy products and detection of milk dilution (Chapter 5). The results show that the ultrasonic technique could be used for compositional characterisation of dairy products, and has the potential to detect diluted milk.

Using the milk authentication method developed in Chapter 4, milk fraud occurrence in the Chinese market was determined, and its relationship with the fraud vulnerability data of the Chinese milk supply chain was elucidated (Chapter 6). Twelve out of 52 tested samples were suspected to be adulterated, of which the protein, fat, or lactose contents were below their respective thresholds. Proportionally more suspect samples were discovered among the products from central-northern and eastern areas in China. Chapter 3 previously also showed higher vulnerability levels for a number of fraud factors in these areas, which indicates an interesting relationship between fraud vulnerability and prevalence.

In conclusion, this thesis highlighted the complexity of food fraud vulnerability in milk supply chains in the Netherlands and China. It also revealed the relationship between vulnerabilities and fraud prevalence as well as the importance of adequate fraud detection systems. It is recommended that a proper guidance for milk fraud mitigation should concern continuous assessment of fraud vulnerabilities in businesses and across the supply chain in combination with availability of efficient and affordable authentication technology.

Acknowledgements

This thesis is the first English book I have written, and this would never have happened without the support of the fantastic people I have met in the past five years.

I first would like to express my sincere gratitude to my promotor Prof. Dr Saskia van Ruth. You encouraged me to step into an emerging project, and your expertise helped me finish this challenge. I truly thank you for your patience in improving my scientific skills, and motivating me to be an independent scientist. In a scientific world where is still male dominated, your female power has encouraged me to go further and firmly as a scientist.

I could not have started the first chapter without the help of my promotor Prof. Dr Wim Huisman. Your professional knowledge of human behaviour was absolutely crucial for initiating the project. I am grateful for your comments for the paperwork, but also for the heart-warming talks which mitigated my concerns regarding the progress of my project.

I would also like to extend my deepest appreciation to my co-promotor Assoc. Prof. Dr Kasper Hettinga. I could always approach you whenever I requested, you could inspire me with professional ideas and never let me down. I am grateful for your constructive scientific advice, but also for the casual talks about the Dutch culture. I was just so lucky to have you as my supervisor.

Special thanks go to my supervisors Assoc. Prof. Dr Liebing Zhang (张列兵) and Dr Jeroen Heck. 特别感谢张老师的知遇之恩。您会在我课题最艰难的时候援手相助，会在我深陷低谷时鼓励指引。从初次见面至今转眼已经十余年，感谢您从未间断的学业上及生活中的支持。Dear Jeroen, our first meeting is dating back to five years ago in Beijing for my first interview. Many thanks for sharing your opinions during our pleasant meetings from the Hutongs in Beijing to the office of FrieslandCampina in Wageningen.

Thanks to this PhD journey, I have met fantastic colleagues, who are much more than colleagues. To Valentina, Isabelle, Alfred, Ningjing, Jing, and Ashraf in WFSR (RIKILT), I will always remember our jokes, trips, chats, and of course scientific discussions in and out of the office.

I gratefully acknowledge Leen van Ginkel and Annemieke Pustjens in WFSR, whose help cannot be overestimated. Thank you for your support and encouragement in the final (but difficult) stage of the PhD. I would also like to express thanks to Vincenzo, thanks for the pleasant chats during the PhD trip, and I am lucky that I was in a group with such an excellent chair.

I gratefully acknowledge the support from the colleagues from Sino-Dutch Dairy Development Centre (SDDDC). Prof. Dr Tiny van Boekel, Engr Kees de Koning, Prof. Shengli Li, and Prof. Dr Toon van Hooijdonk, I sincerely appreciate your scientific opinions on my

们在荷兰还能常相见。

感谢几位相识 20 多年的老友。谢谢闺蜜平，从喀纳斯到涠洲岛，从芬兰到韩国，感恩童年的遇见，共同的成长，和从未间断的同行。谢谢你和李博士，还有 summer，豆瓣，是你们让我在国内的生活有了归属感。同样感谢张希玥，刘洋，旺，朱，冯总，二哥，走过少年的林林总总和成年的平淡悲喜，感恩我们在回归生活的柴米油盐后还能畅谈，以及你们对一个女博士的宽容。

谢谢我的爸爸妈妈。"幸运的人一生都被童年治愈"，我很感恩拥有完美的父母，以及你们给予的美好童年。谢谢你们理解我辞去稳定的工作，支持我走向未知的远方，相信我能够克服所有困难，并一直做我最坚强的后盾。你们是我此生最大的财富。同时也感谢我的表姑，小姨，姑姑，哥哥，嫂子，妹妹，姐姐，弟弟在我读博期间对家里的支持。

Yuzheng Yang

Overview of completed training activities

Discipline specific activities

Dutch Dairy Chain (2016), WA, Leeuwarden, NL.

Postgraduate Food Fraud Symposium (2016), WFSR, Wageningen, NL.

World Dairy Summit (2016), IDF, Rotterdam, NL.

Sensory Perception and Food Preference: Affective Drivers of Food Choice (2016), VLAG, Wageningen, NL.

Advanced Food Analysis (2017), VLAG, Wageningen, NL.

China Dairy Research System Annual Meeting (2017), National Dairy Industry Technology System, Beijing, CN.

5th International Symposium on Dairy Cow Nutrition and Milk Quality (2017), CAAS, Beijing, CN.

5th Duxes China Dairy Summit (2017), Duxes, Beijing, CN.

China Dairy Exhibition (2017), Dairy association of China, Nanjing, CN.

World Dairy Summit (2017), IDF, Belfast, UK.

Food Fraud Conference (2018), New Food, London, UK.

ASSET (2018), QUB, Belfast, UK.

Chemometrics (2018), VLAG, Wageningen, NL.

Postgraduate Food Fraud Symposium (2018), Prague, CZ.

Healthy Food Design (2018), VLAG, Wageningen, NL.

Dairy Protein Biochemistry (2018), VLAG, Wageningen, NL.

Healthy and Sustainable Diets: Synergies and Trade-offs (2019), VLAG, Wageningen, NL.

33rd EFFoST International Conference (2019), Rotterdam, NL.

General courses

VLAG PhD week (2016), VLAG, Soest, NL.

Philosophy and Ethics of Food Science and Technology (2016), VLAG, Wageningen, NL.

Social Dutch 1 (2016), Wageningen in'to Languages, Wageningen, NL.

Data Science: Multiple Imputation in Practice (2017), Utrecht, NL.

Data Science: Data Analysis and Visualization (2017), Utrecht, NL.

Scientific Writing (2017), WGS, Wageningen, NL.

Workshop Cross Cultural Communication (2017). WFSR, Wageningen, NL.

Mobilising Scientific Network (2018), WGS, Wageningen, NL.

Career Orientation (2019), WGS, Wageningen, NL.

Other activities

Preparation of research proposal (2015-2016), WFSR, Wageningen, NL.

Expertise group/business unit meetings (2015-2020), WFSR, Wageningen, NL.

Project meetings with stakeholders (2015-2020), WFSR, Wageningen, NL.

Colloquia (2015-2019), FQD, Wageningen, NL.

PhD study trip to Australia (2018), AU.

Dutch Dairy Experience (2019), SDDDC, NL.

Teaching obligations

Mentoring MSc- and BSc-students and internships (2016-2018). WUR, Wageningen, NL.

Food Fraud and Mitigation (teaching assistant), FQD-36306 (2017-2018). FQD, Wageningen, NL.

List of abbreviations

CAAS: Chinese Academy of Agricultural Sciences

FQD: Food Quality and Design

SDDDC: Sino Dutch Dairy Development Centre

VLAG: Graduate School for Nutrition, Food Technology, Agrobiotechnology and Health Sciences

WA: Wageningen Academy

WFSR: Wageningen Food Safety Research

WGS: Wageningen Graduate School

WUR: Wageningen University & Research

About the author

Yuzheng Yang was born on the 13th March 1987 in Tangshan, Hebei, China. She obtained her bachelor's degree in food quality and safety at Hebei University of Science and Technology in 2010, and obtained her master's degree in food science at China Agricultural University in 2012. Her Master thesis focused on the influence of various salts on heat stability of whey protein concentrates, under the supervision of Dr Zhang Liebing. After graduation she worked at the nutrition department in Beijing YouAn hospital. She was responsible for development of quality assurance procedures for catering services, and customised nutritional recipes for inpatients. In October 2015, she started her PhD study on Food fraud vulnerability assessments and product authentication in Dutch and Chinese milk supply chains, at Wageningen University and Research, under the supervision of Prof. Dr Saskia van Ruth, Prof. Dr Wim Huisman, and Dr Kasper Hettinga. The results of the PhD project are presented in this thesis.

List of publications

Yang, Y., Hettinga, K. A., Erasmus, S. W., Pustjens A. M., & van Ruth, S. M. (2020). Opportunities for fraudsters: When would profitable milk adulterations go unnoticed by common, standardized FTIR measurements? *Food Research International*, *136*, 109543.

Yang, Y., Huisman, W., Hettinga, K. A., Liu, N., Heck, J., Schrijver, G. H., Gaiardoni, L., & van Ruth, S. M. (2019). Fraud vulnerability in the Dutch milk supply chain: Assessments of farmers, processors and retailers, *Food Control*, *95*, 308-317.

Yang, Y., Huisman, W., Hettinga, K. A., Zhang, L., & van Ruth, S. M. (2020). The Chinese milk supply chain: A fraud perspective, *Food Control*, *113*, 107211.

Yang, Y., Wright, W. M. D., Hettinga, K. A., & van Ruth, S. M. (2021). Exploration of an ultrasonic pulse echo system for comparison of milks, creams, and their dilutions. *LWT - Food Science and Technology, 136,* 110616.

Yang, Y., Zhang, L., Hettinga, K. A., Erasmus, S. W., & van Ruth, S. M. (2020). Prevalence of milk fraud in the Chinese market and its relationship with fraud vulnerabilities in the chain. *Foods*, *9*, 9060709.

Liu, N., Pustjens, A. M., Erasmus, S. W., **Yang, Y.**, Hettinga, K. A., & van Ruth, S. M. (2020). Dairy farming system markers: The correlation of forage and milk fatty acid profiles from organic, pasture and conventional systems in the Netherlands, *Food Chemistry*, *314*, 126153.

van Ruth, S. M., Luning, P. A., Silvis, I. C. J., **Yang, Y.**, & Huisman, W. (2018). Differences in fraud vulnerability in various food supply chains and their tiers, *Food Control, 84*, 375-381.