Raw milk quality in large-scale farms under hot weather conditions: Learnings from one-year quality control data

Rodney J Feliciano, Géraldine Boué, Fahad Mohssin, Mohammed Mustafa Hussaini, Jeanne-Marie Membré

Journal of Food Composition and Analysis 117 (2023) 105127

Contents lists available at ScienceDirect

Journal of Food Composition and Analysis

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

Raw milk quality in large-scale farms under hot weather conditions: Learnings from one-year quality control data

Rodney J. Feliciano [a,*], Géraldine Boué [a], Fahad Mohssin [b], Mohammed Mustafa Hussaini [b], Jeanne-Marie Membré [a]

[a] Secalim, INRAE, Oniris, 44307 Nantes, France
[b] AlSafi Danone, Al-Kharj, Saudi Arabia

ARTICLE INFO

Keywords:
Food quality
Food safety
Physicochemical properties
Dairy supply chain
Dairy farm
Seasonality
Statistical analysis

ABSTRACT

A study on the quality control data from a large-scale dairy farm located under hot weather conditions was conducted. Physicochemical properties, microbial count data, and environmental variables (i.e., mean temperature and relative humidity) were examined. The analyses performed were Spearman's rank correlation, principal components analysis (PCA), and partial least squares regression (PLS). The correlation analysis revealed individual correlations between similar variables but weak between physicochemical properties and microbial counts. PCA identified low structure within the dataset but interestingly some seasonal patterns. The predictive modelling approach performed through PLS aimed to predict microbial counts, fat, and protein content using the physicochemical and environmental variables. Microbial counts were not well predicted, while the PLS model satisfactorily predicted fat and protein contents. These two physicochemical properties are associated with delivery payments for raw milk. This study characterized and identified relationships between the properties of raw milk. The utility of the statistical tools was demonstrated in understanding the quality control data. The results highlighted the need to consider data beyond the values regularly monitored in raw milk quality. Ultimately, these results can aid decision-making to improve raw milk quality.

1. Introduction

Raw milk is the precursor of all dairy products, and its properties influence the quality of its derived products. The common raw milk quality values monitored are compositional properties, somatic cell counts (SCC), microbial counts, drug residues, and off-flavors (Murphy et al., 2016). These are associated with the payment schemes between farms and processors. Deviation of the delivered batch from the stipulated raw milk quality values will lead to a penalty or premium payments (Murphy et al., 2016). However, maintaining raw milk quality is a challenge, given the influence exerted by environmental conditions. In particular hot weather conditions have been linked to a decrease in milk yield, changes in physicochemical properties (Kouřimská et al., 2014; Maciuc et al., 2017; Zhou et al., 2013), and microbial diversity (Cempírková, 2007; Li et al., 2018; Rios-Muñiz et al., 2019). Furthermore, raw milk has been associated with increased presence of *Escherichia coli, Staphylococcus aureus,* and thermophilic *Campylobacter* during the warmer summer months (Bertasi et al., 2016; Fairbrother and Nadeau, 2006; Lan et al., 2017). Moreover, the dairy supply chain is reported to face additional constraints due to climate change through a decline in milk yield and an increase of mastitis on farms (Feliciano et al., 2020; Guzmán-Luna et al., 2022; Kim et al., 2022; Mauger et al., 2015).

Statistical tools, on the other hand, have been employed in analysing raw milk quality control data. These tools were used to understand the relationships between raw milk quality parameters (Hanuš et al., 2010; Risoluti et al., 2020), seasonal variation (El-Tahawy and El-Far, 2010; Najafi et al., 2009), analysis of trace elements (Vojnovic and Procida, 1991), and detection of adulteration (Bassbasi et al., 2014; Nikolaou et al., 2020). A similar approach was adopted by Risoluti et al. (2020) in developing a MicroNIR screening platform based on a PLS model that can identify the source (e.g., cow, goat, etc.) and milk quality. Prediction

* Correspondence to: Secalim, INRAE, Oniris, Site de la Chantrerie, CS 40706, 44307, Nantes Cédex 3, France.
E-mail address: rodney.feliciano@oniris-nantes.fr (R.J. Feliciano).

https://doi.org/10.1016/j.jfca.2023.105127
Received 6 July 2022; Received in revised form 28 November 2022; Accepted 3 January 2023
Available online 5 January 2023

of raw milk quality traits was recently demonstrated by Frizzarin et al. (2021), who advocated the use of statistical methods based on machine learning (e.g., neural networks). Despite these, relatively few similar studies were applied in large-scale dairy farms, especially those located in hot weather conditions.

Large-scale dairy farming is a type of intensive farming in which a large number of cows, i.e., more than 1000 cows, are located in a single geographical area. These farms operate more efficiently and are characterized by a higher level of mechanization than conventional farms. Furthermore, they have measures (e.g., cooling systems) that mitigate the effects of severe weather conditions, such as excessive year-round temperatures. As a result, dairy farming is possible in hot-weather regions like the Middle East, which sustains a dairy industry that contributes to the economy despite the unfavorable conditions (Alqaisi et al., 2010). Also, these dairy farms have been known to be among the top milk producers in the world. Nevertheless, large-scale dairy farms operating in hot weather conditions indicate the challenges to be expected in the near future for other farms that will face the effects of climate change.

In this context, the objective of this study was to analyse quantitatively a one-year quality control dataset obtained from a large-scale farm in the Middle East. To this end, the use of the dataset in combination with several statistical tools was deployed. The correlation between the physicochemical properties, microbial counts of raw milk, and climate variables (i.e., daily temperature and humidity) was investigated. The trends and dataset structure were analysed using correlation analysis, principal component analysis (PCA), and partial least-squares (PLS) regression. Analysis of data from a large farm operating in hot weather conditions may help stakeholders in other parts of the world to understand what might happen in their farms under climate change.

2. Materials and methods

2.1. Data collection

The data used for this study were based on one-year raw milk quality control data from a large-scale dairy farm. This dairy farm is situated in a single geographical area in Al-Kharj, Saudi Arabia, and grouped into eight locations (Groups a-h). No special treatment exists between the locations which merely reflect the subdivision of the large dairy farm for implementing farm management practices. The largest of these locations is Group C which contains the greatest number of cows. This dairy farm is mainly dedicated to milk production from the Holstein breed, and its operation follows established animal health management and quality control protocols. The quality control datasets were initially made of separate files of bulk milk colony counts and physicochemical analysis. The former has 1696 data points from colony counts of total aerobic bacteria, thermophiles, and *E. coli,* and the latter contains 5949 points of physicochemical data. One data point corresponds to one sample taken for analysis of milking at each milk production per day. For physicochemical analysis, the samples taken range from 2 to 8 samples per day, while for the microbial counts, 1–3 samples were taken per day. The difference between the number of sample points is that milk is not produced every day for some locations due to differences in calving and days in milk.

The two data files were combined into a single matrix by extracting similar data points sampled on the same collection date and location within the farm. Data points without similar dates and locations were discarded. This combination resulted in a matrix containing 860 data points. Subsequently, the columns of environmental variables and classification into four seasons were added based on the day of sample collection. The former is based on the mean daily temperature and relative humidity of Al Kharj city (Virtualcorssing.com). The collection dates were categorized based on the seasons in Saudi Arabia (visitsaudi.com): winter (December to February), spring (March to May), summer (June to August), and autumn (September to November). Finally, this integrated matrix comprises 19 columns and 473 rows after removing rows with incomplete values (Table 1). The largest sampling location contains 265 points (Group c), followed by 46 points (Group a), and the remaining six subgroups with 27 points each.

The experimental procedures for determining microbial counts and physicochemical properties are briefly explained below. The total aerobic counts were determined by diluting 20 μL of raw milk in 20 mL of distilled water. 1 mL sample of the diluted raw milk was then pour plated with 20 mL of cooled milk agar (Oxoid, Ltd., UK). The plates were then incubated for 48 hrs at 25–30 °C. Ultimately, these were enumerated after 48 h using an automatic colony counter (Flash & Go, IUL Instruments, Spain). The *E. coli* analyses were performed according to NF ISO 4832 (updated in 2006). Briefly, an undiluted 1 mL of raw milk sample was transferred into Petri dishes, followed by the addition of 12–15 mL cooled (45 ± 1 °C) Violet Red Bile Agar (VRBA) (Oxoid, Ltd., UK). This basal-sample medium was allowed to solidify, followed by adding an overlay of VRBA (1–2 mL). It was then allowed to solidify and incubated at 37 ± 1 °C for 48 h. The colonies that presented pink or purplish red colonies with a reddish zone of precipitated bile (≥0.5 mm diameter) were counted using the colony counter. The determination of total thermophilic counts (LPC) was performed by heating 10 mL of raw milk sample at 64 °C for 35 min to eliminate non-thermoduric bacteria. Subsequently, 1 mL of the cooled, heated milk was diluted in 10 mL sterilized distilled water and plated in milk agar (Oxoid, Ltd., UK), and incubated at 25–30 °C for 48 h. Colonies were enumerated automatically using the colony counter. Ultimately, the microbial counts were converted into log values before statistical analyses. Meanwhile, the physicochemical properties were obtained using FTIR Spectrophotometer, MilkoScan FT-120 (Foss A/S, Ltd., Denmark). These properties comprise percent protein, fat, lactose, urea, free fatty acid, total solids, solids not fat, urea, freezing point depression, titratable acidity, and citric acid.

2.2. Statistical analyses

2.2.1. Correlation analyses

The relationship between the physicochemical properties, microbial counts, and environmental variables was determined through correlation analysis using the Spearman rank method. This method allowed the comparison of direct or indirect relationships between variables. The statistical analyses were performed and implemented using R software (R Core Team, 2020). The correlation matrix was constructed using the Spearman method using the built-in cor function and was plotted using the corrplot package (Wei and Simko, 2021). Significant correlations between variables were visualized in the correlation plot at $P > 0.05$ significance level. The blue color denotes positive correlation values, and red is for negative values.

2.2.2. Principal component analysis (PCA)

The PCA was performed to determine the underlying relationship between physicochemical, microbial, and environmental data (i.e., relative humidity and mean temperature) that could not be seen directly by correlation analysis. The dataset within the previously integrated matrix was subjected to PCA with auto-scaling as data pre-treatment. In addition, this dataset was analysed according to the eight sampling locations (Groups a to h) on the farm. Aside from identifying the relationships between the variables, the seasonal patterns were checked in each location on the dairy farm. Subsequently, a location on the farm presenting a visible seasonal pattern was selected. This location is then used to identify the key variables associated with each season. The PCA analyses were performed in R software (R Core Team, 2020) using the

Table 1
Integrated data matrix of physicochemical properties, microbial counts and environmental variables. The presented values are a subset of the complete dataset.

		Physico-chemical properties											Microbial counts (Log CFU/mL)			Environmental variables		
Date of Collection	Location in the farm	Density (g/ cm^3)	Protein (%)	Fat (%)	Total Solids (%)	Solids-Not- Fat (%)	Lactose (%)	Freezing Point Depression (°C)	Titratable Acidity (%)	Citric Acid (%)	Urea (%)	Free Fatty Acid (%)	Total aerobic counts	Lab pasteurized counts	*E. coli* counts	Mean Temperature	Relative Humidity	Season
04/03/2019	A	1032	3.05	3.43	12.22	8.77	4.86	0.54	6.74	0.15	0.04	4.39	4.00	2.80	1.70	17.2	22.89	Spr
08/07/2019	A	1031	2.98	3.7	12.43	8.71	4.86	0.53	6.52	0.15	0.04	4.69	4.46	2.63	1.60	38.7	9.64	Sum
01/10/2019	A	1032	3.12	3.29	12.09	8.82	4.82	0.53	6.74	0.16	0.04	4.28	3.30	2.68	1.62	39.6	14.41	Aut
17/12/2019	A	1032	3.17	3.5	12.4	8.91	4.87	0.54	6.14	0.15	0.04	3.61	3.78	1.95	2.08	16.5	72.44	Wint
01/04/2019	B	1031	3.14	3.5	12.34	8.82	4.79	0.54	6.94	0.15	0.04	4.71	4.34	2.95	2.20	27.4	24.20	Spr
24/06/2019	B	1032	3.22	3.29	12.22	8.88	4.83	0.54	7.05	0.14	0.04	3.97	4.20	3.60	1.73	35.0	12.09	Sum
16/09/2019	B	1032	3.25	3.51	12.59	9.08	4.93	0.56	7.35	0.16	0.04	5.11	4.15	3.31	1.08	38.2	11.85	Aut
23/12/2019	B	1033	3.4	3.64	12.77	9.14	4.85	0.56	7.35	0.14	0.04	4.03	4.04	3.18	2.23	16.1	24.25	Wint
27/05/2019	C	1031	3.18	3.3	12.19	8.85	4.81	0.54	6.94	0.14	0.04	4.33	4.11	2.43	1.60	33.9	15.92	Spr
27/08/2019	C	1032	3.21	3.45	12.35	8.91	4.81	0.53	6.78	0.15	0.05	3.54	4.18	2.32	1.41	36.2	10.51	Sum
26/11/2019	C	1032	3.03	3.33	12.13	8.84	4.95	0.53	6.02	0.15	0.04	4.9	4.08	2.46	1.30	22.4	54.19	Aut
08/01/2019	C	1032	3.25	3.2	12.27	9.03	4.88	0.53	7.00	0.15	0.04	4.88	4.18	3.20	1.62	17.40	50.69	Wint
29/04/2019	D	1031	2.9	3.44	12.14	8.69	4.92	0.54	6.10	0.16	0.04	2.96	3.00	1.48	1.00	27.5	30.40	Spr
22/07/2019	D	1032	2.94	3.08	11.74	8.62	4.87	0.53	6.50	0.14	0.04	4.39	3.00	1.78	0.90	39.4	9.54	Sum
11/11/2019	D	1032	3.07	3.86	12.69	8.86	4.89	0.54	6.64	0.15	0.04	5.09	3.30	2.41	0.90	24.0	52.32	Aut
04/02/2019	D	1031	3.17	4.02	12.87	8.91	4.79	0.55	6.69	0.16	0.04	4.42	3.90	1.78	2.08	15.3	40.84	Wint
13/05/2019	E	1032	2.94	3.52	12.16	8.67	4.85	0.52	6.27	0.15	0.04	3.43	4.23	2.45	1.48	31.3	26.60	Spr
10/06/2019	E	1031	2.9	3.78	12.38	8.53	4.84	0.53	6.09	0.14	0.04	5.06	5.16	3.61	2.08	39.2	9.80	Sum
25/11/2019	E	1033	3.21	3.18	12.21	9.02	4.92	0.55	7.16	0.14	0.04	5.08	4.61	2.26	1.70	18.3	70.74	Aut
23/12/2019	E	1033	3.39	3.39	12.56	9.18	4.88	0.55	7.07	0.15	0.05	3.63	5.10	3.11	1.98	16.1	24.25	Wint
18/02/2019	F	1031	3.15	3.59	12.38	8.7	4.71	0.54	7.03	0.14	0.04	5.78	3.48	3.23	1.60	16.1	27.04	Wint
04/03/2019	F	1031	3.14	3.56	12.4	8.81	4.84	0.54	6.97	0.14	0.04	3.37	4.30	3.02	2.29	17.2	22.89	Spr
24/06/2019	F	1032	2.92	2.87	11.57	8.66	4.9	0.53	6.19	0.15	0.04	4.29	4.08	2.36	1.85	35.0	12.09	Sum
02/09/2019	F	1032	3.26	3.27	12.33	9.05	4.89	0.55	7.47	0.15	0.04	4.25	4.04	2.90	2.41	39.4	10.02	Aut
18/03/2019	G	1032	2.89	2.91	11.45	8.54	4.84	0.52	5.64	0.13	0.04	2.48	4.45	2.11	2.43	19.5	21.22	Spr
24/06/2019	G	1032	3.27	3.3	12.26	8.90	4.82	0.55	7.06	0.13	0.04	3.11	3.70	2.15	1.20	35.0	12.09	Sum
30/09/2019	G	1032	3.17	3.19	12.11	8.94	4.87	0.54	6.98	0.15	0.04	4.96	4.30	3.30	1.48	39.3	13.32	Aut
23/12/2019	G	1033	3.39	3.36	12.53	9.19	4.91	0.55	6.85	0.15	0.04	4.47	3.70	3.08	3.05	16.1	24.25	Wint
04/03/2019	H	1031	3.09	3.72	12.48	8.75	4.85	0.54	6.55	0.13	0.04	4.76	3.00	2.30	1.60	17.2	22.89	Spr
24/06/2019	H	1032	3.03	3.1	11.85	8.69	4.88	0.53	6.46	0.13	0.04	3.55	4.30	3.09	1.56	35	12.09	Sum
02/09/2019	H	1032	3.02	3.45	12.26	8.8	4.94	0.54	6.68	0.15	0.04	4.53	4.04	2.32	1.58	39.4	10.02	Aut
18/02/2019	H	1031	2.95	3.44	11.97	8.51	4.76	0.52	6.07	0.13	0.04	3.63	3.70	1.60	1.70	16.1	27.04	Wint

PCA function of the FactoMineR package (Lê et al., 2008). The plots of the PCA analysis were also generated using the factoextra package developed by Kassambra and Mundt (2020).

2.2.3. Partial least squares (PLS)

The partial least squares regression was performed on Group C, which contains the highest number of cows. The PLS was performed using two modelling approaches. The first is to estimate a model based on the physicochemical properties to predict the total aerobic counts. The second modelling approach was to predict the protein and fat content using the physicochemical properties and environmental data. The PLS analysis was done by fitting the model using the orthogonal scores algorithm (NIPALS), while the optimum number of components included in the predictive model was selected with the lowest root mean squared error of the prediction (RMSEP), reflecting the smallest prediction error. The validation step to determine the quality of the PLS model was performed using the Leave-one-out procedure. These analyses were implemented in R software using the pls package (Mevik and Wehrens, 2007).

3. Results

3.1. Analysis of correlation between microbial counts, physicochemical properties, and environmental variables

The results of the Spearman correlation analysis are presented for the three variable types of raw milk (Fig. 1). The correlation analysis has shown that the microbial counts were correlated with each other (0.29–0.39). Notably, total aerobic counts (logtbc) have a positive correlation with thermophiles (loglpc) at 0.35 and *E. coli* (logcoli) at 0.39. Comparing different variables, microbial counts had a weak positive correlation with most physicochemical properties. An example is for total aerobic counts (logtbc) with percent protein (protein.pt) at 0.14 and percent solids not fat (SNF.pt) at 0.15. In terms of environmental variables, the three microbial counts had a non-significant correlation with mean temperature and humidity, except thermophilic counts, which had a positive correlation with humidity (0.10).

Meanwhile, the physicochemical properties had mostly positive correlation values with each other (0.15–0.92). An example is percent protein with total solids (TS) at 0.61, titratable acidity (TA) at 0.84, and freezing point depression at 0.80. The physicochemical variables were compared with environmental variables and were found to contain some weak negative correlation. An example is mean temperature with fat (−0.28) and total solids (−0.21). These reflect the negative impact of temperature on raw milk and can be linked with its biological impact on cow lactation. However, mean temperature has a neutral effect on protein content, while humidity had a relatively weak correlation with percent citric acid (−0.12), percent urea (−0.16), and titratable acidity (−0.09). These instances can be linked with the effectivity of current practices in limiting the impact of high temperatures. Overall, these results have shown interesting one-at-a-time correlations across the different variables. However, a similar correlation with multiple variables can be beneficial if looked into. Also, determining the underlying data structure between variables may be helpful in understanding the current dataset. Pattern analysis, such as principal components analysis (PCA), can determine these underlying data structure.

3.2. Results of the principal components analysis

The principal component analyses were performed by analysing the raw milk datasets according to their location within the farm. The first two principal components (PCs) were estimated at 25.4–33.9 % (PC 1) and 14.5–27 % (PC 2) across the different locations, while the cumulative variance of these two were 44.31–59.25 % (Table 2). These values mean that the PCA could not detect a high data structure in the datasets and thus, the 3rd and 4th PCs were also checked. The physicochemical properties were mostly positively correlated with the 1st and 2nd PCs across the different datasets within the farm (Supplementary Table S1). However, the microbial counts had more occasions of a negative correlation with the 2 PCs. The values ranged for total aerobic counts (−0.09 to 0.49), thermophiles (−0.07 to −0.13), and *E. coli* (−0.19 to −0.53). On the other hand, the mean temperature is more positively correlated with the 3rd and 4th PCs (−0.29 to 0.74), while humidity is more positively related with the 1st and 2nd PCs (−0.51 to 0.61) across the different locations within the farm.

Despite having a low data structure detected within the dataset, seasonal clusters were seen by plotting the individual data points. These were emphasized across the eight different locations on the farm (Fig. 2a-h). These clusters can be primarily linked with the environmental variables, namely, mean temperature and humidity during raw milk collection (Fig. 3a-b). These two variables were observed to have an inverse relationship. A minimal variability was observed for the mean temperature and relative humidity during summer compared to other seasons (e.g. autumn). Meanwhile, overlaps were observed for some data points collected during winter-spring and autumn-summer. However, in Group C the overlaps of points were seen across the four seasons. These can be linked with the variability in the physicochemical properties.

Furthermore, to understand the seasonal patterns, the dataset of Group F was analysed and was found to be characterized by PC 1 and PC 2 (Table 3). Key variables that contributed the most to the development of the respective PCs were identified. PC 1 which captured the 27.5 % of the variance in the data was positively described by the physicochemical parameters, namely, solids not fat (0.95), protein (0.91), density (0.63), titratable acidity (0.62), and urea (0.60). Meanwhile, PC 2, had both positive and negative correlations for density (−0.58), lactose (−0.61), fat (0.90), total solids (0.86), and urea (−0.31). In contrast, the *E.coli* counts of raw milk weakly contributed to the seasonal pattern observed

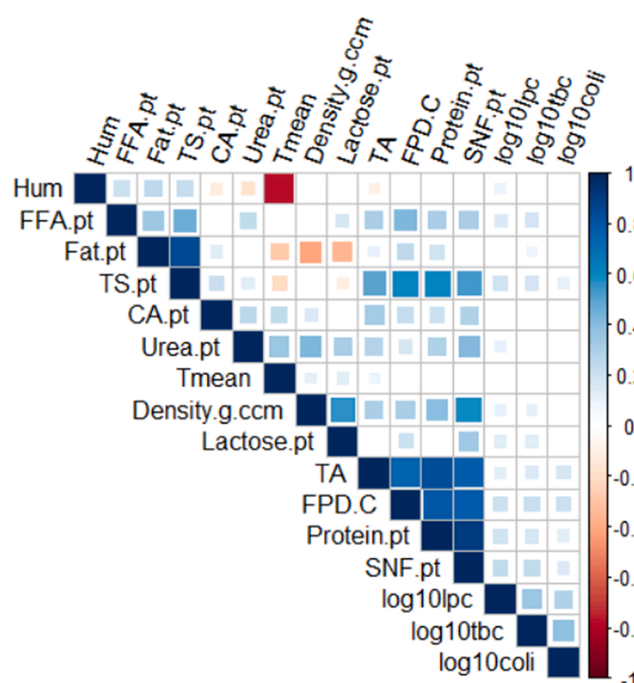

Fig. 1. Spearman rank correlation of the different physicochemical properties, microbial counts of raw milk and the environmental conditions during the day these are collected.Legend: % concentration of: free fatty acid (FFA.pt), fat (fat.pt), protein (Protein.pt), lactose (Lactose.pt), total solids (TS.pt), citric acid (CA.pt), urea (Urea.pt) and others such as density (Density.g.cm), titratable acidity (TA), freezing point depression (FPD.C), solids not fat (SNF.pt), total aerobic counts (logtbc), lab pasteurized counts (log10lpc), total coliforms (log10coli), mean temperature (Tmean), and humidity (Hum).

in this location (-0.14). A quick overview of these key physicochemical properties shows that these were relatively stable across the year despite the changes in mean temperature and humidity (Fig. 3c-j). This relative stability on the properties of the raw milk can be linked with the effectivity of the current dairy farming operation in minimizing the influence of extreme weather conditions in the dairy farm. Moreover, the utility of these variables in predicting other variables of raw milk quality (e.g., microbial count) can be further explored.

3.3. Partial least squares regression

The partial least squares regression analyses were performed on the largest dataset from location C of the farm. The goal was to explore further the utility of the physicochemical properties and environmental data in predicting the other relevant properties of raw milk. Therefore, two modelling attempts were performed by predicting microbial counts and the physicochemical properties (fat and protein content) in raw milk.

3.3.1. Attempt to predict microbial counts using physicochemical properties and environmental variables

The first attempt constructed a model to predict the microbial counts in raw milk using the regularly monitored physicochemical properties of raw milk and the two environmental variables in dairy farms during milking. This technique would have given food safety managers on the farm a tool to gauge the food safety status even before the standard plate count methods are released. However, this resulted with a model that was able to capture a low percentage of variance for *E. coli* (15.31 %), thermophiles (11.69 %), and total aerobic counts (15.31 %). The predictive efficiency was also reflected in the R^2 and root mean square error of prediction (RMSEP) plots, which are both measures of fit statistics (Fig. 4). First, in the R^2 plots it was shown that the values have a very low correlation even if the number of components in the model is increased. Second, the RMSEP plots show consistently high values even if the maximum number of components is included.

Ultimately, these show that the physicochemical properties in raw milk cannot be used to predict all the microbial counts. Moreover, this does not mean that correlation does not exist rather linear regression models and linear-based estimation methods were not able to capture the predictive potential underlying the data structure. Nevertheless, these are similar to the results of correlation analysis, where a weak positive correlation was observed for these two variables. In addition, these may be linked with the results of the PCA where the first two PCs could explain only half of the variance of the data.

3.3.2. Attempt to predict % fat and % protein using other physicochemical variables

The second attempt was performed to construct two separate PLS models to determine the predictability of fat and protein content in milk using the other remaining physicochemical properties (namely, density, total solids, solids not fat, lactose, freezing point depression, titratable acidity, urea, citric acid, and free fatty acid) with the two environmental variables. The prediction of these two properties is significant since both have an implication on milk delivery payments where deviations in the agreed fat and protein content will lead to penalties on the delivered batch. The results of these models show that the physicochemical properties and environmental variables could predict both the percent protein and fat, as shown by RMSEP and R^2 plots.

Furthermore, the model developed was simplified by selecting the optimum number of components or latent variables. This was done through cross-validation using the Leave-one-out method, where the optimal number of components was selected by examining the RMSEP plots where the lowest value was observed at 6 and 7 components for predicting fat and protein, respectively (Fig. 5a-b). The adjusted CV obtained with six optimal components were at 0.06 (fat content) and 0.03 (protein content), while the variance explained during the training phase were at 94.11 % and 95.18 %, respectively. In addition, the R^2 was also shown to reach a high value as these components are incorporated in the model (Fig. 5c-d). The final models predicting fat and protein content are presented in the 2 equations below.

$$\% \text{ Fat} = 101{\cdot}14 + 0{\cdot}85 \text{ (Total Solids)} - 0{\cdot}17 \text{ (Titratable acidity)} - 0{\cdot}10 \text{ (Density)} - 0{\cdot}05 \text{ (Lactose)} + 0{\cdot}03 \text{ (Solids Not Fat)} + 0{\cdot}01 \text{ (Citric acid)} \tag{1}$$

$$\% \text{ Protein} = 23{\cdot}55 + 0{\cdot}91 \text{ (Solids not fat)} - 0{\cdot}52 \text{ (Lactose)} + 0{\cdot}06 \text{ (Freezing point depression)} - 0{\cdot}04 \text{ (Citric acid)} - 0{\cdot}03 \text{ (Density)} - 0{\cdot}02 \text{ (\%Total Solids)} - 0{\cdot}02 \text{ (Urea)} \tag{2}$$

Another way to look at these models is by analysing the influence that the variables exert in predicting the results. This can be seen with the values of the estimated coefficients of the PLS model. For determining the % fat, the variable that contributed most is total solids (0.85) followed by titratable acidity (−0.17), density (−0.10), lactose (−0.05), solids not fat (0.03), and (0.01) citric acid. With these results, the influence of solids not fat, density, and total solids in determining fat, are not necessarily new. However, the influence of titratable acidity is worth taking into consideration. On the other hand, it is interesting to observe that the prediction of percent protein is influenced negatively by lactose (−0.53) and citric acid (−0.04), while the influence of solids not fat (0.91), freezing point depression (0.06), and density (−0.05) are expected as these are influenced by the protein content in raw milk. In fact, these were measured to assure that raw milk is not adulterated with water and that the quality of the batch was up to standard.

4. Discussion

The current study evaluated the interrelationships between the physicochemical properties, microbial counts, and environmental variables. Several insights can be gained through these correlation results. The current microbial concentration levels not affecting the physicochemical properties of the raw milk as seen from their weak correlation. In contrast with those observed by Yuan et al. (2022), where the correlation between the two means that the microorganisms found in raw milk are already modifying its physicochemical properties (e.g. *Acinetobacter* with TA). As such, microbial counts should be kept at current levels in order to ensure that these will not have negative effects on the physicochemical properties of future raw milk deliveries.

The TBC counts are mostly LPC and *E. coli* counts as shown by their relative high correlation with each other. Comparison with the literature shows that these are similar to Jayarao et al. (2004) for TBC with LPC (0.51) or *E. coli* (0.39) and Pantoja et al. (2009) for TBC with *E. coli* (0.41) and LPC (0.17). Farm management practices can be directed towards the control of *E. coli* by revisiting maintaining animal health (e.g. low microbial counts in cow feed) and farm practices (e.g. change in bedding materials) (Fairbrother and Nadeau, 2006; LeJeune and Kauffman, 2005). LPC counts can be further reduced by revisiting hygiene protocols in frequent contamination sources such as milking equipment (Jindal et al., 2016). The weak correlation between the LPC and *E. coli* counts shows their non-relation and are similar to the ones

Table 2
First five principal components estimated from the datasets per farm groups in the dairy farm.

Group	%Variance				
	PC 1	PC 2	PC 3	PC 4	PC 5
A	30.35	21.95	12.99	8.27	7.47
B	33.85	22.20	10.72	7.54	6.64
C	29.80	14.52	12.85	9.23	6.16
D	27.45	23.43	13.20	10.82	7.79
E	33.34	17.79	13.80	8.97	6.15
F	25.35	22.39	13.15	9.41	8.60
G	33.76	23.01	10.01	7.59	6.71
H	32.24	27.01	10.29	7.14	6.41

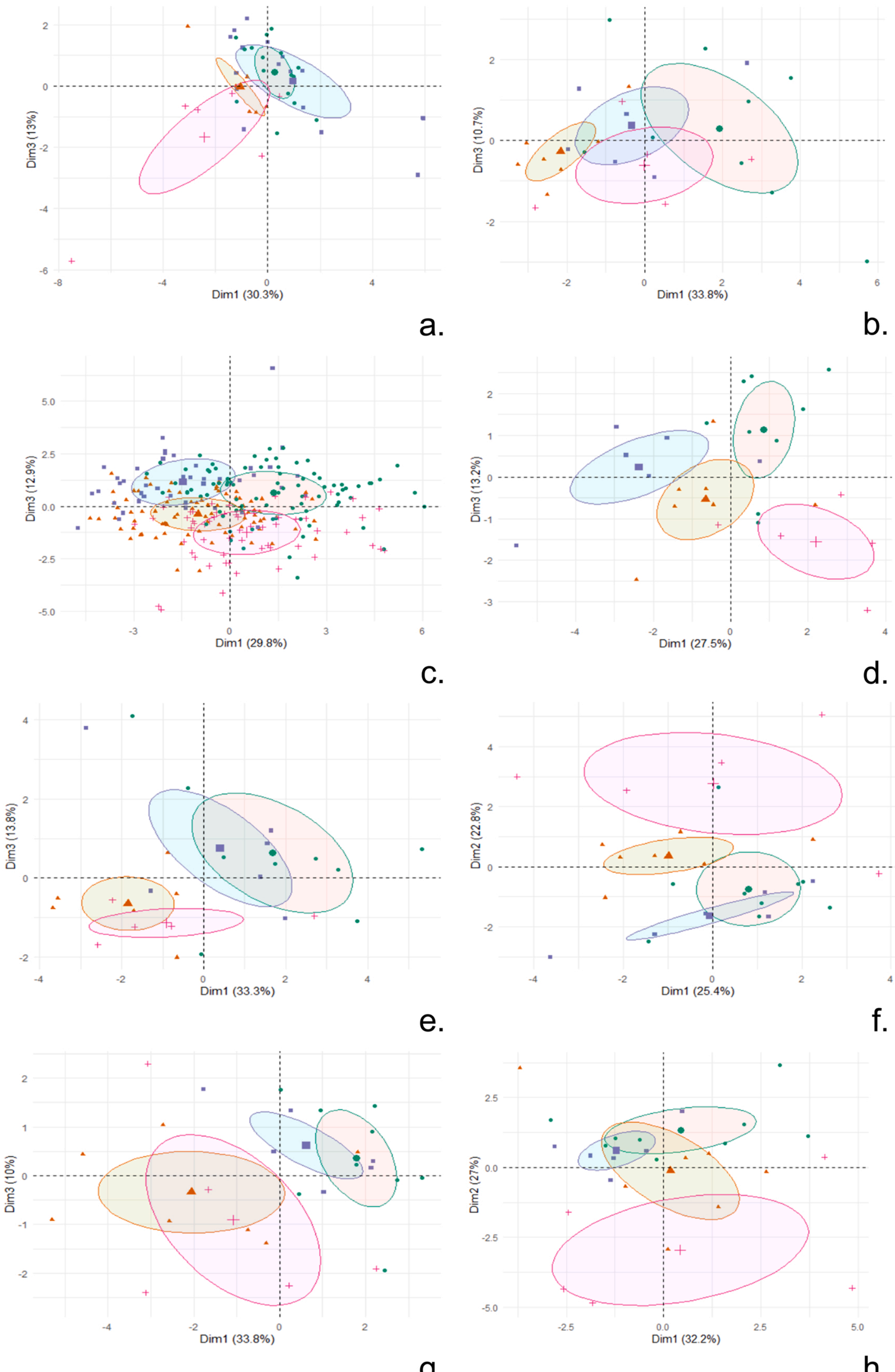

Fig. 2. PCA plot of the individual data points from the 8 different locations (Groups a-h) in the dairy farm showing varying degrees of the seasonal clusters. The seasons are autumn (●), spring (▲), summer (■), and winter (+).

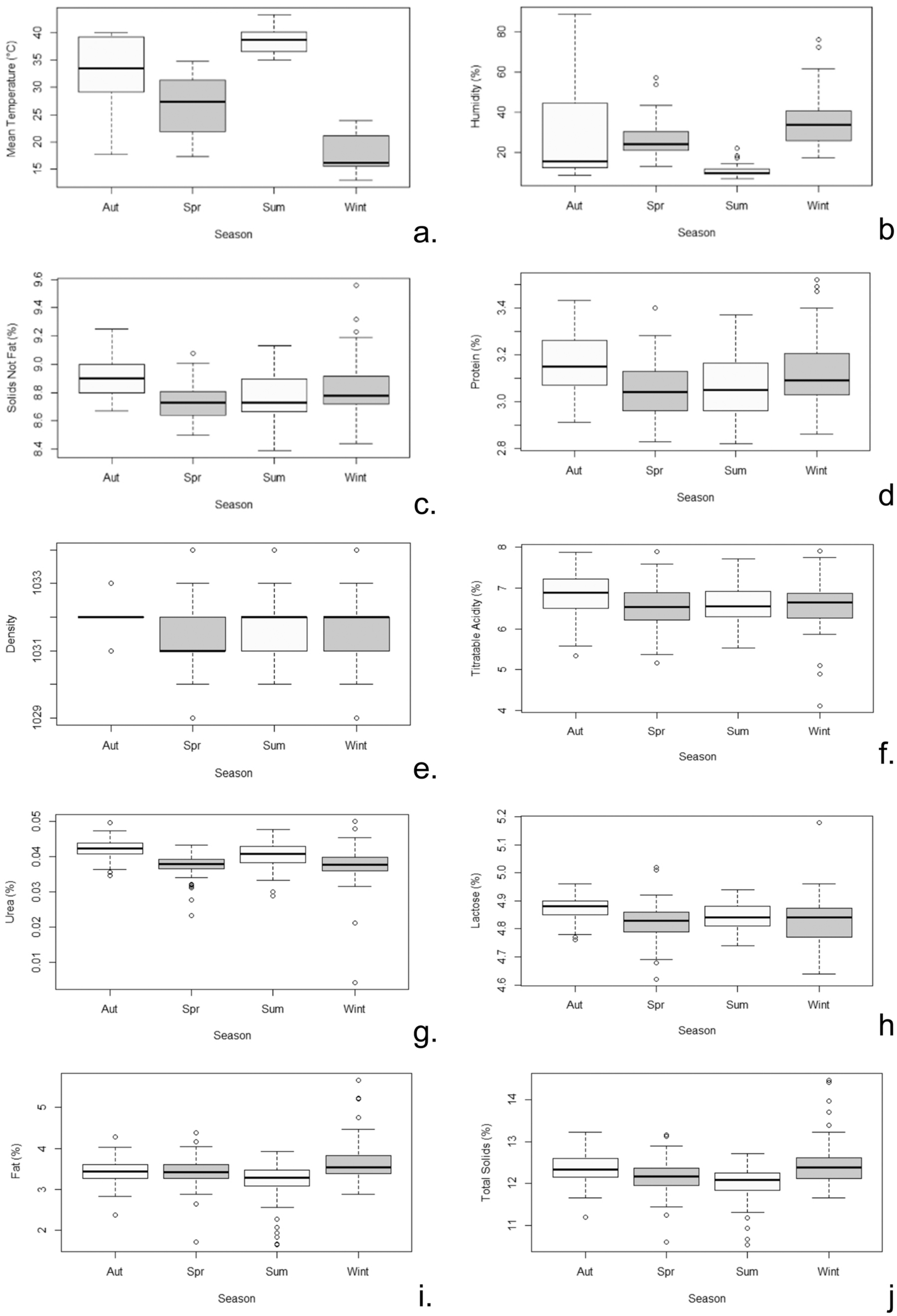

Fig. 3. Boxplot per season of key variables identified through PCA. a. mean temperature, b. humidity, c. solids not fat, d. protein, e. density, f. titratable acidity, g. total aerobic counts, h. lactose, i. fat, j. total solids.

Table 3
Principal component loadings estimated for group F.

	PC 1	PC 2	PC 3	PC 4
Density.g.ccm	0.63	-0.58	-0.01	0.13
Protein.pt	0.91	0.24	-0.18	0.10
Fat.pt	0.05	0.90	0.26	-0.22
TS.pt	0.31	0.86	0.26	-0.17
SNF.pt	0.95	0.03	0.18	0.06
Lactose.pt	0.12	-0.61	0.73	0.07
FPD.C	0.91	0.06	0.17	0.20
TA	0.62	0.04	-0.54	0.21
CA.pt	0.21	-0.26	0.44	-0.72
Urea.pt	0.60	-0.31	-0.01	-0.30
FFA.pt	0.04	0.19	-0.62	0.01
log10tbc	-0.10	-0.39	-0.10	-0.08
log10lpc	0.16	0.21	0.47	0.53
log10coli	-0.38	-0.14	0.38	0.61
Tmean	0.13	-0.73	0.02	-0.19
Hum	0.00	0.61	0.32	-0.08

*The red color denotes negative values while green denotes the positive values greater than 0.50.

reported by Jayarao et al. (2004) at (0.08) and Pantoja et al. (2009) at (0.17). These variations in the microorganisms in raw milk can be associated with on-farm practices (Elmoslemany et al., 2010) or bacterial shedding occurrences during high-temperature summer conditions (Fairbrother and Nadeau, 2006). Ultimately, these show what hazards need to be controlled in the latter part of the supply chain (Pantoja et al., 2009).

Meanwhile, the correlation between physicochemical properties reflects the association between the main constituents of milk and its properties (e.g., protein content with freezing point depression and total solids). These correlations between the physicochemical properties are similar to the reports of Karlsson et al. (2017) for protein with fat and total solids determined through the Pearson correlation method. Furthermore, in terms of the correlation with mean temperature on physicochemical properties and microbial counts, these seem to be limited. Nevertheless, farm management practices must direct attention to temperature sensitive physicochemical properties such as the protein and fat content (Bernabucci et al., 2014; Yang et al., 2013).

The PCA has shown some seasonal patterns across the dairy farm despite the relatively low data structure. Nevertheless, the estimated PCs are comparable with those in the literature (Chen et al., 2014; Priyashantha et al., 2021). Chen et al. (2014) reported PC 1 (37.79 %) and PC 2 (17.22 %), for raw milk obtained in the UK. Similarly, Priyashantha et al. (2021) have reported low PC 1 (20–41 %) and PC 2 (6–15 %) during the selection of influential farm factors for raw milk properties obtained in Northern Sweden. Whereas, Bassbasi et al. (2014) have reported higher PCs (PC1:91% and PC2: 8 %) in their respective datasets.

In this study, key physicochemical properties were also identified and characterized throughout the year. This can particularly be seen with the main constituents of raw milk, namely, protein (3.05–3.16 %), fat (3.22–3.62 %), and lactose (4.82–4.88 %) content. However, comparison with the literature varies depending on the study. An example is the higher reported values from raw milk obtained in Italy for protein (3.29–3.50 %), lactose (5.15–5.30 %), and fat (3.20–3.80 %) (Bernabucci et al., 2015). Similarly, for the values obtained from farms in Northern Sweden in terms of protein (3.44–3.46 %), fat (4.14–4.19%), and lactose (4.71–4.73%) (Karlsson et al., 2017). On the other hand, the current values are closer to milk samples obtained in Iran in terms of protein (3.12–3.15 %) and fat (3.51–3.63 %) (Najafi et al., 2009). This is also similar to those reported by Bassbasi et al. (2014) for raw milk in Morocco obtained during cold and hot periods for protein content (2.96–3.30 %) and fat (3.19–4.19 %). Another example is the milk obtained from Friesian cows in Egypt, where the protein (2.29–3.23 %), fat (3.08–3.88 %), and lactose (3.53–4.88 %) were lower compared to this farm (El-Tahawy and El-Far, 2010). Furthermore, these values also vary when compared with the milk composition from other cow breeds such as those obtained from Brown Swiss cows in Italy at 3.71–3.81 (% protein) and 4.31–4.52 (% fat) (Bittante et al., 2015). Nevertheless, these values are still within the range reported in the literature for protein (2.3–4.4 %), fat (2.5–5.5 %), and lactose content (3.8–5.3 %) (Walstra et al., 2005). Overall, the relative stability of the physicochemical properties throughout the year can also be linked with the attenuated influence of environmental variables on dairy milk production. This means that current dairy farming operations effectively mitigate the effects of environmental conditions. Furthermore, the low variance explained by the estimated PCs points to an opportunity to be exploited by revisiting other datasets collected within the farm. In terms of food quality management, this underscores the need to go beyond the common data kept by the quality control department. These may include differences in practices throughout the seasons (e.g. feed type, bedding, heat stress). This is in line with the insights presented in the literature

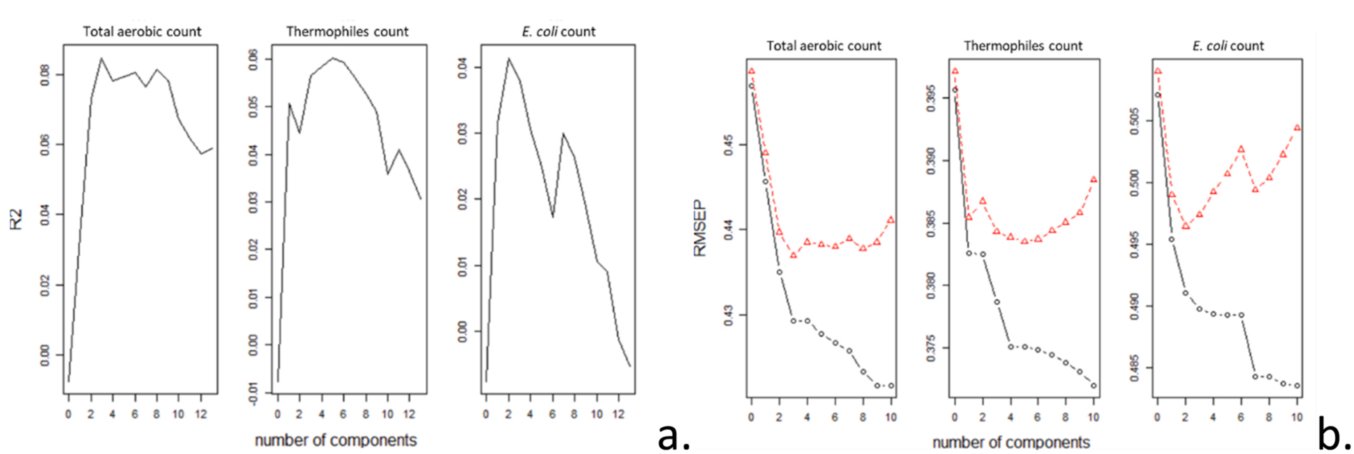

Fig. 4. The predictive efficiency of the PLS model in predicting the microbial counts in raw milk. a. R^2 plots and b. RMSEP plots. For the RMSEP plots the black line denotes the calibration run while the red line denotes the Leave-one-out. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

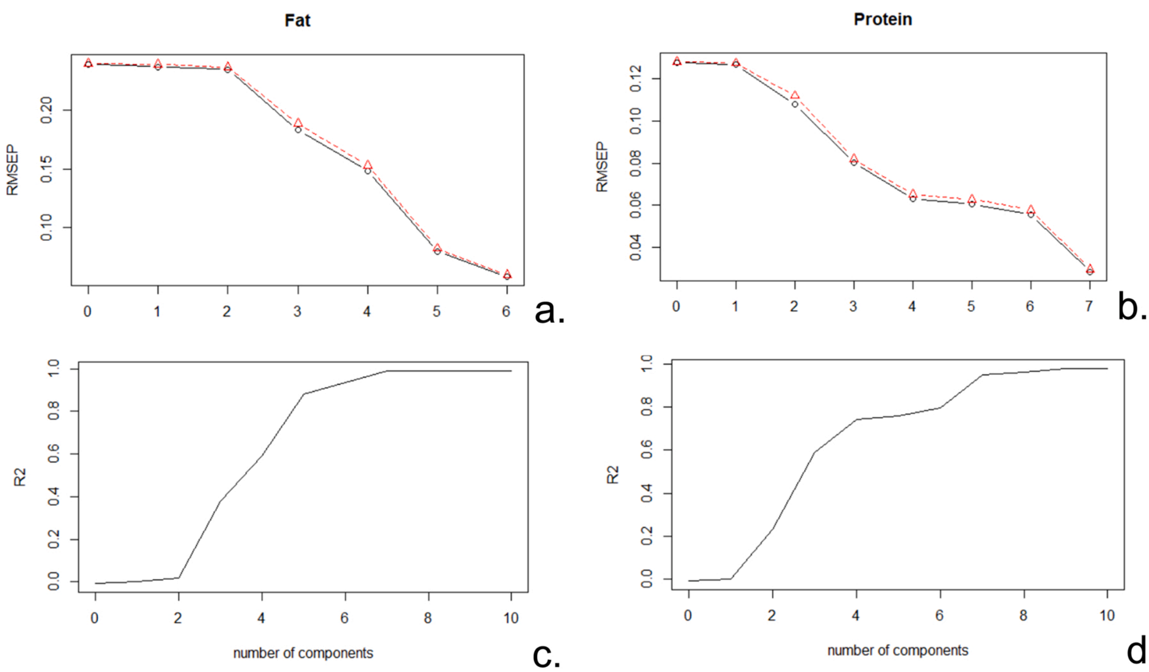

Fig. 5. The predictive efficiency of the PLS model in predicting the % Fat and % Protein in raw milk as shown by the RMSEP plots (a-b), and R^2 plots (c-d). For the RMSEP plots the black line denotes the calibration run while the red line denotes the Leave-one-out. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

where the influence of on-farm practices (milking system, housing) and feeds during specific seasons on raw milk have been underscored (Bassbasi et al., 2014; Priyashantha et al., 2021).

Meanwhile, the PLS modelling approach used in this study demonstrated the possibility of predicting raw milk properties. This approach is in contrast with the literature where PLS is used in analysing the milk spectra to predict the protein content and the functionality of milk samples (e.g. cheese making) (Bittante et al., 2021; El Jabri et al., 2020). On the other hand, the same approach was not able to predict microbial counts. This is due to several factors that influence its contamination and its ability to survive given the intrinsic raw milk properties. Furthermore, this reflects the results of the low correlation results between variable types and the low contribution of microbial counts in describing the estimated PCs. As such, this supports the need to check other datasets that can be used in explaining and reducing microbial counts. The same can also be said about the prediction of percentage fat and protein, where there are other datasets that can further explain these properties.

5. Conclusion

The current study has shown three different ways of analysing the quality control dataset of raw milk. The correlation analysis has revealed the association between the same variable types but not with microbial counts and physicochemical properties. Meanwhile, the PCA has shown a low relative structure within the dataset but some degree of seasonal patterns were observed. In turn, these cannot be explained by temperature and relative humidity across the season alone. The modelling approach using PLS regression was done through the physicochemical properties, temperature, and relative humidity as predictors. This resulted in the prediction of percentage fat and protein but not microbial content.

In general, these analyses highlight the relevance of physicochemical properties in describing the datasets. Continuous monitoring of these and the possible use of the model when monitoring farm operations can benefit decision-making practices. However, the relatively low contribution of microbial counts in describing the data and model prediction point to the need for a different approach to analysing these data. This can be done by analysing other datasets that may influence microbial counts (e.g., environmental data). Quality control departments can benefit from this comparison with data that is not usually associated with raw milk quality. Ultimately, the current paper highlights the utility of these data analysis tools in understanding the quality control data in other dairy farms to aid in improving the quality of raw milk produced.

The farm evaluated in this study can be considered as representative of a large-scale dairy farm in hot weather conditions. The results revealed low data structure within the dataset and the relative stability of key variables year-round. These can be attributed to the effectivity of pre-requisite programmes, i.e. good agricultural practices and good hygiene practices, in the farm analysed in our study. To prepare for climate change, other potentially vulnerable dairy farms may need to review their pre-requisite programmes and even to consider more stringent food safety mitigation strategies.

CRediT authorship contribution statement

Rodney J. Feliciano: Data curation, Methodology, Formal analysis, Writing – review & editing. **Géraldine Boué:** Supervision, Visualization. **Fahad Mohssin:** Visualization, Investigation. **Mohammed Mustafa Hussaini:** Visualization, Investigation. **Jeanne-Marie Membré:** Conceptualization, Methodology, Formal analysis, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgements

This research has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie-ITN grant agreement No. 813329, (PROTECT) (http://www.protect-itn.eu/).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jfca.2023.105127.

References

Alqaisi, O., Ndambi, O.A., Uddin, M.M., Hemme, T., 2010. Current situation and the development of the dairy industry in Jordan, Saudi Arabia, and Syria. Trop. Anim. Health Prod. 42, 1063–1071. https://doi.org/10.1007/s11250-010-9553-y.

Bassbasi, M., Platikanov, S., Tauler, R., Oussama, A., 2014. FTIR-ATR determination of solid non fat (SNF) in raw milk using PLS and SVM chemometric methods. Food Chem. 146, 250–254. https://doi.org/10.1016/j.foodchem.2013.09.044.

Bernabucci, U., Basiricò, L., Morera, P., Dipasquale, D., Vitali, A., Piccioli Cappelli, F., Calamari, L., 2015. Effect of summer season on milk protein fractions in Holstein cows. J. Dairy Sci. 98, 1815–1827. https://doi.org/10.3168/jds.2014-8788.

Bernabucci, U., Biffani, S., Buggiotti, L., Vitali, A., Lacetera, N., Nardone, A., 2014. The effects of heat stress in Italian Holstein dairy cattle. J. Dairy Sci. 97, 471–486. https://doi.org/10.3168/jds.2013-6611.

Bertasi, B., Losio, M.N., Daminelli, P., Finazzi, G., Serraino, A., Piva, S., Giacometti, F., Massella, E., Ostanello, F., 2016. Seasonal variability of thermophilic *Campylobacter* spp. in raw milk sold by automatic vending machines in Lombardy. Ital. J. Food Saf. 5, 2014–2016. https://doi.org/10.4081/ijfs.2016.5848.

Bittante, G., Cecchinato, A., Tagliapietra, F., Schiavon, S., Toledo-Alvarado, H., 2021. Effects of breed, farm intensiveness, and cow productivity level on cheese-making ability predicted using infrared spectral data at the population level. J. Dairy Sci. 104, 11790–11806. https://doi.org/10.3168/jds.2021-20499.

Bittante, G., Cipolat-Gotet, C., Malchiodi, F., Sturaro, E., Tagliapietra, F., Schiavon, S., Cecchinato, A., 2015. Effect of dairy farming system, herd, season, parity, and days in milk on modeling of the coagulation, curd firming, and syneresis of bovine milk. J. Dairy Sci. 98, 2759–2774. https://doi.org/10.3168/jds.2014-8909.

Cempírková, R., 2007. Contamination of cow's raw milk by psychrotrophic and mesophilic microflora in relation to selected factors. Czech J. Anim. Sci. 52, 387–393. https://doi.org/10.17221/2325-cjas.

Chen, B., Lewis, M.J., Grandison, A.S., 2014. Effect of seasonal variation on the composition and properties of raw milk destined for processing in the UK. Food Chem. 158, 216–223. https://doi.org/10.1016/j.foodchem.2014.02.118.

El-Tahawy, A.S., El-Far, A.H., 2010. Influences of somatic cell count on milk composition and dairy farm profitability. Int. J. Dairy Technol. 63, 463–469. https://doi.org/10.1111/j.1471-0307.2010.00597.x.

El Jabri, M., Trossat, P., Wolf, V., Beuvier, E., Rolet-Répécaud, O., Gavoye, S., Gaüzère, Y., Belysheva, O., Gaudillière, N., Notz, E., Grosperrin, P., Laithier, C., Delacroix-Buchet, A., 2020. Mid-infrared spectrometry prediction of the cheese-making properties of raw Montbéliarde milks from herds and cheese dairy vats used for the production of Protected Designation of Origin and Protected Geographical Indication cheeses in Franche-Comté. J. Dairy Sci. 103, 5992–6002. https://doi.org/10.3168/jds.2019-17491.

Elmoslemany, A.M., Keefe, G.P., Dohoo, I.R., Wichtel, J.J., Stryhn, H., Dingwell, R.T., 2010. The association between bulk tank milk analysis for raw milk quality and on-farm management practices. Prev. Vet. Med. 95, 32–40. https://doi.org/10.1016/j.prevetmed.2010.03.007.

Fairbrother, J.M., Nadeau, E., 2006. Contaminación de animales por *Escherichia coli* en la finca. Rev. Sci. Tech. l'OIE 25, 555–569. https://doi.org/10.20506/rst.25.2.1682.

Feliciano, R.J., Boué, G., Membré, J.-M., 2020. Overview of the potential impacts of climate change on the microbial safety of the dairy industry. Foods 9, 1794. https://doi.org/10.3390/foods9121794.

Frizzarin, M., Gormley, I.C., Berry, D.P., Murphy, T.B., Casa, A., Lynch, A., McParland, S., 2021. Predicting cow milk quality traits from routinely available milk spectra using statistical machine learning methods. J. Dairy Sci. 104, 7438–7447. https://doi.org/10.3168/jds.2020-19576.

Guzmán-Luna, P., Nag, R., Martínez, I., Mauricio-Iglesias, M., Hospido, A., Cummins, E., 2022. Quantifying current and future raw milk losses due to bovine mastitis on European dairy farms under climate change scenarios. Sci. Total Environ. 833 https://doi.org/10.1016/j.scitotenv.2022.155149.

Hanuš, O., Frelich, J., Tomáška, M., Vylětelova, M., Genčurová, V., Kučera, J., Třináctý, J., 2010. The analysis of relationships between chemical composition, physical, technological and health indicators and freezing point in raw cow milk. Czech J. Anim. Sci. 55, 11–29. https://doi.org/10.17221/1708-cjas.

Jayarao, B.M., Pillai, S.R., Sawant, A.A., Wolfgang, D.R., Hegde, N.V., 2004. Guidelines for monitoring bulk tank milk somatic cell and bacterial counts. J. Dairy Sci. 87, 3561–3573. https://doi.org/10.3168/jds.S0022-0302(04)73493-1.

Jindal, S., Anand, S., Huang, K., Goddard, J., Metzger, L., Amamcharla, J., 2016. Evaluation of modified stainless steel surfaces targeted to reduce biofilm formation by common milk sporeformers. J. Dairy Sci. 99, 9502–9513. https://doi.org/10.3168/jds.2016-11395.

Karlsson, M.A., Langton, M., Innings, F., Wikström, M., Lundh, Å.S., 2017. Short communication: Variation in the composition and properties of Swedish raw milk for ultra-high-temperature processing. J. Dairy Sci. 100, 2582–2590. https://doi.org/10.3168/jds.2016-12185.

Kassambra, A., Mundt, F., 2020. Factoextra: Extract and Visualize the Results of Multivariate Data Analyses. ⟨https://cran.r-project.org/web/packages/factoextra/index.html⟩.

Kim, S.H., Ramos, S.C., Valencia, R.A., Cho, Y.Il, Lee, S.S., 2022. Heat stress: effects on Rumen microbes and host physiology, and strategies to alleviate the negative impacts on lactating dairy cows. Front. Microbiol. 13, 1–23. https://doi.org/10.3389/fmicb.2022.804562.

Kouřimská, L., Legarová, V., Panovská, Z., Pánek, J., 2014. Quality of cows' milk from organic and conventional farming. Czech J. Food Sci. 32, 398–405. https://doi.org/10.17221/510/2012-cjfs.

Lan, X.Y., Zhao, S.G., Zheng, N., Li, S.L., Zhang, Y.D., Liu, H.M., McKillip, J., Wang, J.Q., 2017. Short communication: microbiological quality of raw cow milk and its association with herd management practices in Northern China. J. Dairy Sci. 100, 4294–4299. https://doi.org/10.3168/jds.2016-11631.

Lê, S., Josse, J., Husson, F., 2008. FactoMineR: an R package for multivariate analysis. J. Stat. Softw. 25, 1–18. https://doi.org/10.18637/jss.v025.i01.

LeJeune, J.T., Kauffman, M.D., 2005. Effect of sand and sawdust bedding materials on the fecal prevalence of Escherichia coli O157:H7 in dairy cows. Appl. Environ. Microbiol. 71, 326–330. https://doi.org/10.1128/AEM.71.1.326-330.2005.

Li, N., Wang, Y., You, C., Ren, J., Chen, W., Zheng, H., Liu, Z., 2018. Variation in raw milk microbiota throughout 12 months and the impact of weather conditions. Sci. Rep. 8, 1–10. https://doi.org/10.1038/s41598-018-20862-8.

Maciuc, V., Gabriela Radu-rusu, C., Cristian Popescu, E., Mihail Radu-rusu, R., Claudiu Jurco, E., 2017. Influence of season and cows farming system on milk physical, chemical and hygienic traits. Rom. Biotechnol. Lett. 22, 13108–13119. https://www.rombio.eu/vol22nr6/10_Maciuc_et_al.pdf.

Mauger, G., Bauman, Y., Nennich, T., Salathé, E., 2015. Impacts of climate change on milk production in the United States. Prof. Geogr. 67, 121–131. https://doi.org/10.1080/00330124.2014.921017.

Mevik, B.-H., Wehrens, R., 2007. The pls Package: principal component and partial least squares regression in R. J. Stat. Softw. 18, 1–23. https://doi.org/10.18637/jss.v018.i02.

Murphy, S.C., Martin, N.H., Barbano, D.M., Wiedmann, M., 2016. Influence of raw milk quality on processed dairy products: how do raw milk quality test results relate to product quality and yield? J. Dairy Sci. 99, 10128–10149. https://doi.org/10.3168/jds.2016-11172.

Najafi, M., Mortazavi, S.A., Koocheki, A., Khorami, J., Rekik, B., 2009. Fat and protein contents, acidity and somatic cell counts in bulk milk of Holstein cows in the Khorasan Razavi Province, Iran. Int. J. Dairy Technol. 62, 19–26. https://doi.org/10.1111/j.1471-0307.2008.00451.x.

Nikolaou, P., Deskoulidis, E., Topoglidis, E., Kakoulidou, A.T., Tsopelas, F., 2020. Application of chemometrics for detection and modeling of adulteration of fresh cow milk with reconstituted skim milk powder using voltammetric fingerpriting on a graphite/ SiO2 hybrid electrode. Talanta 206, 120223. https://doi.org/10.1016/j.talanta.2019.120223.

Pantoja, J.C.F., Reinemann, D., Ruegg, P.L., 2009. Associations among milk quality indicators in raw bulk milk. J. Dairy Sci. 92, 4978–4987. https://doi.org/10.3168/jds.2009-2329.

Priyashantha, H., Lundh, Å., Höjer, A., Bernes, G., Nilsson, D., Hetta, M., Saedén, K.H., Gustafsson, A.H., Johansson, M., 2021. Composition and properties of bovine milk: a study from dairy farms in northern Sweden; Part I. Effect of dairy farming system. J. Dairy Sci. 104, 8582–8594. https://doi.org/10.3168/jds.2020-19650.

R Core Team, 2020. R: a Language and Environment For Statistical Computing, Vienna, Austria. ⟨https://www.r-project.org/⟩.

Rios-Muñiz, D., Cerna-Cortes, J.F., Lopez-Saucedo, C., Angeles-Morales, E., Bobadilla-Del Valle, M., Ponce-De Leon, A., Estrada-Garcia, T., 2019. Longitudinal analysis of the microbiological quality of raw cow's milk samples collected from three small family dairy farms in Mexico over a 2-year period. J. Food Prot. 82, 2194–2200. https://doi.org/10.4315/0362-028X.JFP-19-155.

Risoluti, R., Gullifa, G., Materazi, S., 2020. Assessing the quality of milk using a multicomponent analytical platform MicroNIR/Chemometric. Front. Chem. 8, 1–8. https://doi.org/10.3389/fchem.2020.614718.

Vojnovic, D., Procida, G., 1991. Chemometric differentiation of raw and commercial milk by trace elements using principal component analysis. Food Addit. Contam. 8, 343–349. https://doi.org/10.1080/02652039109373983.

Walstra, P., Walstra, Pieter, Wouters, J.T.M., Geurts, T.J., 2005. Dairy Science and Technology. CRC Press. https://doi.org/10.1201/9781420028010.

Wei, T., Simko, V., 2021. R Package "corrplot": Visualization of a Correlation Matrix. ⟨https://cran.r-project.org/web/packages/corrplot/index.html⟩.

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. J. Dairy Sci. 96, 6863–6869. https://doi.org/10.3168/jds.2013-6846.

Yuan, H., Han, S., Zhang, S., Xue, Y., Zhang, Y., Lu, H., Wang, S., 2022. Microbial properties of raw milk throughout the year and their relationships to quality parameters. Foods 11, 3077. https://doi.org/10.3390/foods11193077.

Zhou, Y., Li, N., Dong, G., Wu, W., 2013. Impact assessment of recent climate change on rice yields in the Heilongjiang reclamation area of North-East China. J. Sci. Food Agric. 93, 2698–2706. https://doi.org/10.1002/jsfa.6087.