

## Article

# Predictions of Milk Fatty Acid Contents by Mid-Infrared Spectroscopy in Chinese Holstein Cows

Xiuxin Zhao <sup>1,2,†</sup>, Yuetong Song <sup>1,3,4,†</sup>, Yuanpei Zhang <sup>2</sup>, Gaozhan Cai <sup>2</sup>, Guanghui Xue <sup>2</sup>, Yan Liu <sup>2</sup>, Kewei Chen <sup>4</sup>, Fan Zhang <sup>1</sup>, Kun Wang <sup>1,4</sup>, Miao Zhang <sup>1,4</sup> , Yundong Gao <sup>1</sup>, Dongxiao Sun <sup>3,\*</sup> , Xiao Wang <sup>1,\*</sup>  and Jianbin Li <sup>1,\*</sup>

<sup>1</sup> Institute of Animal Science and Veterinary Medicine, Shandong Academy of Agricultural Sciences, Jinan 250100, China

<sup>2</sup> Shandong OX Livestock Breeding Co., Ltd., Jinan 250100, China

<sup>3</sup> Department of Animal Genetics and Breeding, College of Animal Science and Technology, China Agricultural University, Beijing 100193, China

<sup>4</sup> Yantai Institute, China Agricultural University, Yantai 264670, China

\* Correspondence: sundx@cau.edu.cn (D.S.); xiaowangzntc@163.com (X.W.); msdljb@163.com (J.L.)

† These authors contributed equally to this work.

**Abstract:** Genetic improvement of milk fatty acid content traits in dairy cattle is of great significance. However, chromatography-based methods to measure milk fatty acid content have several disadvantages. Thus, quick and accurate predictions of various milk fatty acid contents based on the mid-infrared spectrum (MIRS) from dairy herd improvement (DHI) data are essential and meaningful to expand the amount of phenotypic data available. In this study, 24 kinds of milk fatty acid concentrations were measured from the milk samples of 336 Holstein cows in Shandong Province, China, using the gas chromatography (GC) technique, which simultaneously produced MIRS values for the prediction of fatty acids. After quantification by the GC technique, milk fatty acid contents expressed as g/100 g of milk (milk-basis) and g/100 g of fat (fat-basis) were processed by five spectral pre-processing algorithms: first-order derivative (DER1), second-order derivative (DER2), multiple scattering correction (MSC), standard normal transform (SNV), and Savitzky–Golay convolution smoothing (SG), and four regression models: random forest regression (RFR), partial least square regression (PLSR), least absolute shrinkage and selection operator regression (LassoR), and ridge regression (RidgeR). Two ranges of wavebands (4000–400  $\text{cm}^{-1}$  and 3017–2823  $\text{cm}^{-1}$ /1805–1734  $\text{cm}^{-1}$ ) were also used in the above analysis. The prediction accuracy was evaluated using a 10-fold cross validation procedure, with the ratio of the training set and the test set as 3:1, where the determination coefficient ( $R^2$ ) and residual predictive deviation (RPD) were used for evaluations. The results showed that 17 out of 31 milk fatty acids were accurately predicted using MIRS, with RPD values higher than 2 and  $R^2$  values higher than 0.75. In addition, 16 out of 31 fatty acids were accurately predicted by RFR, indicating that the ensemble learning model potentially resulted in a higher prediction accuracy. Meanwhile, DER1, DER2 and SG pre-processing algorithms led to high prediction accuracy for most fatty acids. In summary, these results imply that the application of MIRS to predict the fatty acid contents of milk is feasible.

**Keywords:** prediction; milk; fatty acid content; mid-infrared spectroscopy; Chinese Holstein cow



**Citation:** Zhao, X.; Song, Y.; Zhang, Y.; Cai, G.; Xue, G.; Liu, Y.; Chen, K.; Zhang, F.; Wang, K.; Zhang, M.; et al. Predictions of Milk Fatty Acid Contents by Mid-Infrared Spectroscopy in Chinese Holstein Cows. *Molecules* **2023**, *28*, 666. <https://doi.org/10.3390/molecules28020666>

Academic Editors: Pamela Manzi and Mena Ritota

Received: 3 December 2022

Revised: 1 January 2023

Accepted: 4 January 2023

Published: 9 January 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Lipids in milk provide a major source of energy and the essential structural components for the cell membranes of the newborns in all mammalian species. They also confer distinctive properties to dairy foods that affect further processing procedures [1]. Milk fat is rich in many fatty acids that are important to human health [2–4]. Studies have shown that more than 400 different fatty acids have been identified in milk fat, but most of them only appeared in trace amounts [5], where around 12 kinds of fatty acids in bovine milk fat

presented at above a 1% concentration [6]. Moreover, changes in milk fatty acids may also affect cow health and energy statuses [7].

Currently, several techniques have been developed to measure fatty acids in milk, including high performance liquid chromatography (HPLC), gas chromatography (GC), near-infrared spectroscopy (NIRS), mid-infrared spectrum (MIRS), etc. [8–10]. Chemical methods (e.g., HPLC and GC) provide high measurement accuracy for fatty acid contents of bovine milk, but their pretreatments are multifarious and costly, causing difficulties in realizing the high-throughput measurements [11,12]. Of note, infrared spectroscopy-based measurement methods show advantages of providing rapid and low-cost predictions of milk fatty acid contents [13]; thus, they have become the promising technologies for high-throughput measurements, but they still need to be optimized to improve their prediction accuracy.

The utilization of infrared spectroscopy to predict the milk fatty acid contents in dairy cattle has been reported in many studies. Coppa et al. (2010) established a prediction equation for milk fatty acid contents based on the NIRS from 468 milk samples that predicted the total milk fatty acids, SFA, MUFA, PUFA, C18:1, and conjugated linoleic acid (CLA), with  $R^2$  values greater than 0.88. Soyeurt et al. (2006) developed a fatty acid prediction model using 600 milk samples from 275 cows of 6 breeds to predict C10:0, C12:0, C14:0, C16:0, C16:1cis-9, C18:1, C18:2cis-9, SFA (saturated fatty acids), and MUFA (monounsaturated fatty acids), based on MIRS data, with the cross-validated coefficients of determination ( $R^2$ ) of 0.62 ~ 0.94. Subsequently, Soyeurt et al. (2011) investigated the MIRS prediction of fatty acids across various cattle breeds, production systems, and countries. They summarized that the usefulness of the built equations providing the best prediction accuracy for animal breeding and milk payment systems was  $R^2 \geq 0.75$  and 0.95, respectively [4]. For the Dutch cattle breeds (Dutch Friesian, Meuse-Rhine-Yssel, Groningen White Headed, and Jersey), Maurice-Van Eijndhoven et al. (2020) updated the calibration equations from the European project RobustMilk [4] using the enlarged datasets and validated their usefulness to predict most milk fatty acids. De Marchi et al. (2011) used 267 milk samples from Brown Swiss cattle to predict fatty acids by MIRS and suggested the implementation of the used prediction models in milk recording schemes on fatty acid contents information for breeding purposes. Fleming et al. (2017) used MIRS to predict fatty acid contents from 373 cows of four breeds and obtained the cross-validation  $R^2$  of 0.60~0.90 for most individual fatty acid models. In addition, the genetic correlations among milk fatty acids predicted by MIRS were also explored in a large-scale milk sampling ( $n = 34,141$ ) of New Zealand dairy cattle, where they implied the application of MIRS as the phenotypic proxy for the genetic selection of fatty acid contents [14]. In the Chinese Holstein population, Du et al. (2020) estimated the heritability of MIRS and several milk production traits, i.e., protein, fat, and lactose percentages, along with their genetic correlations. They found that MIRS heritability ranged from 0 to 0.11 and genetic correlations varied significantly [15]. In sheep, ewes, and goats, MIRS was also used to predict the fatty acid profiles for the establishment and validation of the predictive models [16–18].

Previous studies used a partial least square regression model (non-integrated learning model) [19,20] to investigate the effects of different spectral preprocessing methods on the prediction equation accuracy [4,5,21–23]. However, the combined effects of the regression models and spectral preprocessing methods on the prediction equation accuracy for different fatty acids has rarely been explored, especially for the milk fat of Chinese Holstein cows. Therefore, the objective of this study was to investigate the prediction methods under the optimal strategy to predict milk fatty acids with high accuracy based on the MIRS data from the dairy herd improvement (DHI) database of Chinese Holstein cattle and to potentially provide the high-throughput measurements of a large amount of milk fatty acid phenotypic data; thereby, our study enabled milk fatty acid traits to be feasibly recorded for genetic evaluations of such traits in dairy cattle breeding programs in China. To the best of our knowledge, this is the first time the MIRS predictions on fatty acids of two types of fatty acid measurements (g/100 g of milk and g/100 g fat)

have been investigated with five pre-processed algorithms and two ranges of wavebands (4000~400  $\text{cm}^{-1}$  and 3017~2823  $\text{cm}^{-1}$ /1805~1734  $\text{cm}^{-1}$ ) using four regression models in Chinese Holstein cattle.

## 2. Results

### 2.1. Statistical Description

After quantification by the GC technique, statistical descriptions of individual and grouped fatty acid contents expressed as milk-basis (g/100 g of milk) and fat-basis (g/100 g of fat) are summarized in Table 1. The mean values of the individual fatty acid contents varied from 0.003 (C11:0, C20:1, C20:5n3, and C18:3n6) to 0.877 (C16:0) and their variation coefficients varied from 5.837% (C24:0) to 35.416% (C10:0), when they were expressed as milk-basis (g/100 g of milk). For grouped fatty acid contents, the mean values varied from 0.060 (SCFA) to 1.627 (SFA), and their variation coefficients ranged from 25.514% (PUFA) to 33.392% (SCFA) (Table 1). Similarly, the mean values of individual fatty acid contents varied from 0.094 (C20:5n3) to 28.620 (C16:0), and their variation coefficients varied from 13.802% (C16:0) to 44.207% (C22:1n9), when they were expressed as fat-basis (g/100 g of fat). For grouped fatty acid contents, the mean values varied from 1.934 (SCFA) to 52.710 (SFA), and their variation coefficients ranged from 12.978% (MCFA) to 19.365% (LCFA) (Table 1).

**Table 1.** The minimum, mean, maximum, and variation coefficient of fatty acid contents measured by the GC technique.

Fatty Acid	Milk-Basis (g/100 g of Milk)					Fat-Basis (g/100 g of Fat)				
	Sample Size	Minimum	Mean	Maximum	Variation Coefficient (%)	Sample Size	Minimum	Mean	Maximum	Variation Coefficient (%)
C8:0	325	0.007	0.016	0.028	28.784	324	0.327	0.532	0.757	14.778
C10:0	323	0.013	0.044	0.082	35.416	326	0.598	1.402	2.324	21.273
C11:0	317	0.002	0.003	0.006	25.439	319	0.053	0.110	0.175	23.125
C12:0	321	0.019	0.062	0.115	34.871	327	0.829	2.018	3.323	21.721
C13:0	321	0.003	0.005	0.008	24.468	323	0.082	0.166	0.255	22.062
C14:0	322	0.094	0.231	0.371	27.975	321	4.058	7.546	11.358	15.075
C15:0	320	0.012	0.029	0.052	29.138	321	0.456	0.968	1.519	21.227
C16:0	325	0.366	0.877	1.491	28.154	323	17.710	28.620	42.251	13.802
C17:0	324	0.008	0.016	0.027	26.180	322	0.285	0.528	0.797	19.289
C18:0	320	0.098	0.313	0.600	35.024	326	4.070	10.254	17.150	25.195
C20:0	321	0.006	0.008	0.011	15.098	321	0.157	0.270	0.398	19.278
C22:0	332	0.004	0.005	0.006	9.413	329	0.074	0.172	0.286	25.529
C24:0	324	0.004	0.005	0.005	5.837	323	0.069	0.154	0.246	24.812
C14:1	325	0.006	0.019	0.033	32.157	317	0.230	0.610	1.138	31.033
C16:1	322	0.016	0.038	0.068	30.326	320	0.602	1.258	2.202	26.094
C18:1n9c	321	0.189	0.460	0.815	28.657	323	7.953	15.282	23.746	20.201
C20:1	321	0.003	0.003	0.005	13.134	319	0.065	0.116	0.190	24.333
C22:1n9	322	0.007	0.015	0.028	32.090	317	0.193	0.510	1.140	44.207
C20:3n6	322	0.003	0.006	0.009	24.212	322	0.109	0.192	0.288	18.233
C20:4n6	323	0.004	0.007	0.010	20.368	317	0.122	0.222	0.329	18.416
C20:5n3	332	0.002	0.003	0.004	10.163	323	0.046	0.094	0.149	23.530
C18:2n6c	323	0.030	0.070	0.120	28.437	327	1.052	2.300	3.654	18.149
C18:3n6	321	0.003	0.003	0.004	7.621	318	0.047	0.100	0.169	24.898
C18:3n3	324	0.004	0.008	0.013	24.352	324	0.155	0.278	0.417	16.480
SFA	325	0.692	1.627	2.714	28.415	322	29.494	52.710	74.351	13.287
UFA	323	0.288	0.638	1.090	26.277	326	13.162	21.266	31.904	18.082
MUFA	322	0.240	0.539	0.938	26.878	324	9.501	17.920	27.111	19.321
PUFA	324	0.046	0.098	0.159	25.514	325	1.720	3.196	4.863	16.229
SCFA	323	0.020	0.060	0.109	33.392	324	0.926	1.934	2.936	18.762
MCFA	325	0.580	1.269	2.147	27.475	322	25.617	41.372	61.272	12.978
LCFA	321	0.371	0.925	1.610	28.353	326	17.003	30.776	46.784	19.365

Note: SFA, UFA, MUFA, PUFA, SCFA, MCFA, LCFA, and GC indicate saturated fatty acid, unsaturated fatty acid, monounsaturated fatty acid, polyunsaturated fatty acid, short chain (4 to 10 carbons) fatty acid, medium chain (11 to 16 carbons) fatty acid, long chain (more than 16 carbons) fatty acid, and gas chromatography, respectively. The variation coefficient (%) is the ratio of standard deviation to the mean, which can be used to compare the degree of dispersion among the fatty acids.

## 2.2. Predictions of Milk Fatty Acid Contents

The best prediction accuracy obtained by the optimal strategy from the test set for each fatty acid is summarized in Table 2, after considering different pre-processing algorithms, MIRS ranges, and regression models. In total, 16, 7, 6, and 2 fatty acids achieved the best prediction accuracy when the RFR, LassoR, PLSR, and RidgeR models were used, respectively. Similarly, the DER2, DER1, SG, SNV, and MSC algorithms resulted in 9, 8, 8, 4, and 2 fatty acids for best prediction accuracy, respectively. In addition, 22 fatty acids obtained the best prediction accuracy when they were expressed as g/100 g of milk (milk-basis), but only 9 fatty acids when expressed as g/100 g of fat (fat-basis). For most fatty acids (16/31), the ensemble learning model (RFR), with higher robustness and generalization, produced higher prediction accuracy than those predicted by the non-ensemble learning models (Table 2).

**Table 2.** Best prediction accuracy for the optimal strategy in the test set for each fatty acid.

Fatty Acid	Pre-Processing Algorithm	MIRS Range (cm <sup>-1</sup> )	Model	Basis (g/100 g)	Test Set	
					R <sup>2</sup>	RPD
C8:0	SNV	3017~2823/1805~1734	PLSR	Milk	0.77	2.11
C10:0	DER1	3017~2823/1805~1734	RFR	Milk	0.77	2.07
C11:0	DER1	3017~2823/1805~1734	LassoR	Fat	0.55	1.48
C12:0	DER1	3017~2823/1805~1734	LassoR	Milk	0.84	2.50
C13:0	SG	3017~2823/1805~1734	PLSR	Milk	0.66	1.72
C14:0	DER1	4000~400	RFR	Milk	0.78	2.05
C15:0	SG	3017~2823/1805~1734	PLSR	Milk	0.57	1.53
C16:0	SG	3017~2823/1805~1734	RFR	Milk	0.75	1.98
C17:0	SG	3017~2823/1805~1734	LassoR	Milk	0.73	1.89
C18:0	DER1	4000~400	PLSR	Milk	0.77	2.08
C20:0	SNV	3017~2823/1805~1734	PLSR	Fat	0.82	2.35
C22:0	DER2	4000~400	RFR	Fat	0.86	2.66
C24:0	SG	4000~400	RFR	Fat	0.80	2.20
C14:1	MSC	3017~2823/1805~1734	PLSR	Fat	0.62	1.63
C16:1	SNV	3017~2823/1805~1734	LassoR	Milk	0.62	1.64
C18:1n9c	SG	3017~2823/1805~1734	LassoR	Milk	0.77	2.00
C20:1	DER2	4000~400	RFR	Fat	0.76	2.04
C22:1n9	DER1	4000~400	RFR	Fat	0.65	1.67
C18:2n6c	MSC	4000~400	RFR	Milk	0.63	1.61
C18:3n3	SG	4000~400	RFR	Milk	0.70	1.82
C18:3n6	DER2	3017~2823/1805~1734	RFR	Fat	0.76	2.00
C20:3n6	DER1	4000~400	RFR	Milk	0.62	1.61
C20:4n6	SNV	4000~400	RFR	Milk	0.50	1.42
C20:5n3	DER1	4000~400	RFR	Fat	0.91	3.06
SFA	SG	3017~2823/1805~1734	RFR	Milk	0.76	2.01
UFA	DER2	3017~2823/1805~1734	LassoR	Milk	0.82	2.15
MUFA	DER2	3017~2823/1805~1734	LassoR	Milk	0.79	2.06
PUFA	DER2	4000~400	RidgeR	Milk	0.71	1.75
SCFA	DER2	4000~400	RFR	Milk	0.77	2.04
MCFA	DER2	3017~2823/1805~1734	RFR	Milk	0.75	2.00
LCFA	DER2	3017~2823/1805~1734	RidgeR	Milk	0.83	2.29

Note: MIRS, DER1, DER2, MSC, SNV, SG, RFR, PLSR, LassoR, RidgeR, R<sup>2</sup>, and RPD indicate mid-infrared spectrum, first-order derivative, second-order derivative, multiple scattering correction, standard normal transform, Savitzky–Golay convolution smoothing, random forest regression, partial least square regression, least absolute shrinkage and selection operator regression, ridge regression, determination coefficient, and residual predictive deviation, respectively.

In Table 2, the best prediction accuracy (R<sup>2</sup>) for the optimal strategy showed R<sup>2</sup> values from 0.62 (C20.3n6) to 0.91 (C20.5n3) in the test set for 28 fatty acids, where only 6 fatty acids showed R<sup>2</sup> values higher than 0.8, including C12:0 (0.84), C20:0 (0.82), C22:0 (0.86), C20:5n3 (0.91), UFA (0.82), and LCFA (0.83). For R<sup>2</sup> values higher than 0.75 and RPD values higher than 2, we found 17 fatty acids in total: C8:0 (R<sup>2</sup> = 0.77 and RPD = 2.11),

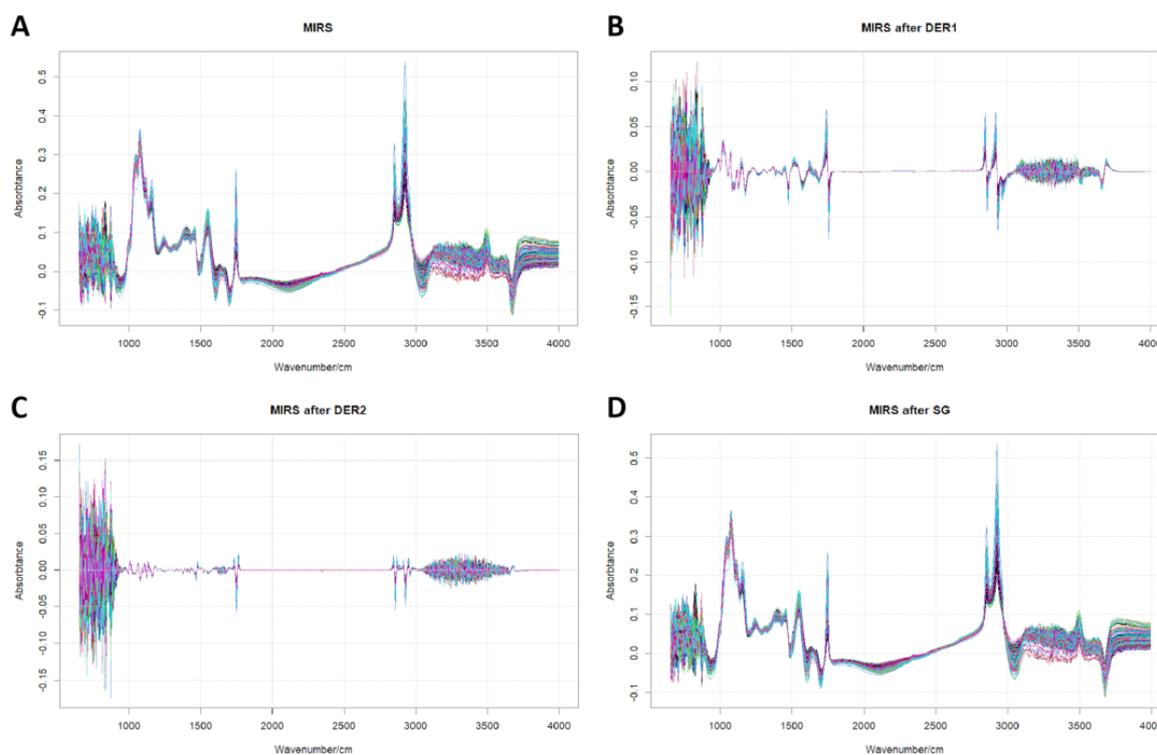
C10:0 ( $R^2 = 0.77$  and RPD = 2.07), C12:0 ( $R^2 = 0.84$  and RPD = 2.50), C14:0 ( $R^2 = 0.78$  and RPD = 2.05), C18:0 ( $R^2 = 0.77$  and RPD = 2.08), C20:0 ( $R^2 = 0.82$  and RPD = 2.35), C22:0 ( $R^2 = 0.86$  and RPD = 2.66), C24:0 ( $R^2 = 0.80$  and RPD = 2.20), C18:1n9c ( $R^2 = 0.77$  and RPD = 2.00), C20:1 ( $R^2 = 0.76$  and RPD = 2.04), C20:5n3 ( $R^2 = 0.91$  and RPD = 3.06), SFA ( $R^2 = 0.76$  and RPD = 2.01), UFA ( $R^2 = 0.82$  and RPD = 2.15), MUFA ( $R^2 = 0.79$  and RPD = 2.06), SCFA ( $R^2 = 0.77$  and RPD = 2.04), MCFA ( $R^2 = 0.75$  and RPD = 2.00), and LCFA ( $R^2 = 0.83$  and RPD = 2.29) (Table 2).

Table 3 shows the best prediction accuracy of different prediction models for each fatty acid, using training and test sets. All prediction accuracies ( $R^2$  and RPD) after four regression model analyses (RFR, PLSR, LassoR, and RidgeR), based on two types of fatty acid measurements (g/100 g of milk and g/100 g fat), two ranges of wavebands (4000~400  $\text{cm}^{-1}$  and 3017~2823  $\text{cm}^{-1}$ /1805~1734  $\text{cm}^{-1}$ ), and five spectral pre-processing algorithms (DER1, DER2, MSC, SNV, and SG), are listed in Supplementary File S1. In the training set, the  $R^2$  values ranged from 0.18 to 0.79, with a mean of 0.58, and RPD values ranging from 1.08 to 2.18, with a mean of 1.59, when expressed as milk-basis (g/100 g of milk). Similarly,  $R^2$  values ranged from 0.13 to 0.90, with a mean of 0.47, and RPD values from 1.07 to 3.20, with mean of 1.52, when expressed as fat-basis (g/100 g of fat). In the test set,  $R^2$  values ranged from 0.14 to 0.84 with mean of 0.66 and RPD values from 1.04 to 2.50 with mean of 1.78 when expressed as milk basis (g/100 g of milk). Similarly, the  $R^2$  values ranged from 0.15 to 0.91 with mean of 0.49 and RPD values from 1.07 to 3.06 with mean of 1.52 when expressed as fat basis (g/100 g of fat) (Table 3). Additionally, the MIRS and processed MIRS after DER1, DER2, and SG pre-processing algorithms are shown in Figure 1.

**Table 3.** Best prediction accuracy of different prediction models for each fatty acid expressed as g/100 g of fat and g/100 g of milk, using training and test sets.

Fatty Acid	Pre-Processing Algorithm		MIRS Range ( $\text{cm}^{-1}$ )		Model		Training Set				Test Set			
	Milk	Fat	Milk	Fat	Milk	Fat	$R^2$		RPD		$R^2$		RPD	
							Milk	Fat	Milk	Fat	Milk	Fat	Milk	Fat
C8:0	SNV	MSC	3017~2823/1805~1734	3017~2823/1805~1734	PLSR	LassoR	0.75	0.43	2.01	1.33	0.77	0.43	2.11	1.32
C10:0	DER1	DER1	3017~2823/1805~1734	3017~2823/1805~1734	RFR	LassoR	0.61	0.49	1.60	1.40	0.77	0.44	2.07	1.33
C11:0	DER2	DER1	3017~2823/1805~1734	3017~2823/1805~1734	LassoR	LassoR	0.57	0.51	1.53	1.43	0.53	0.55	1.46	1.48
C12:0	DER1	SNV	3017~2823/1805~1734	3017~2823/1805~1734	LassoR	LassoR	0.79	0.55	2.18	1.49	0.84	0.27	2.50	1.17
C13:0	SG	SNV	3017~2823/1805~1734	3017~2823/1805~1734	PLSR	LassoR	0.24	0.56	1.16	1.50	0.66	0.42	1.72	1.30
C14:0	DER1	DER1	4000~400	3017~2823/1805~1734	RFR	PLSR	0.66	0.16	1.72	1.10	0.78	0.43	2.05	1.34
C15:0	SG	MSC	3017~2823/1805~1734	3017~2823/1805~1734	PLSR	PLSR	0.45	0.25	1.37	1.17	0.57	0.32	1.53	1.22
C16:0	SG	DER2	3017~2823/1805~1734	4000~400	RFR	RidgeR	0.64	0.55	1.66	1.33	0.75	0.22	1.98	1.12
C17:0	SG	MSC	3017~2823/1805~1734	3017~2823/1805~1734	LassoR	PLSR	0.65	0.40	1.70	1.32	0.73	0.59	1.89	1.56
C18:0	DER1	SNV	4000~400	3017~2823/1805~1734	PLSR	LassoR	0.66	0.60	1.72	1.58	0.77	0.55	2.08	1.49
C20:0	SG	SNV	3017~2823/1805~1734	3017~2823/1805~1734	PLSR	PLSR	0.52	0.76	1.46	2.04	0.71	0.82	1.88	2.35
C22:0	DER2	DER2	4000~400	4000~400	RidgeR	RFR	0.70	0.83	1.76	2.42	0.52	0.86	1.44	2.66
C24:0	DER2	SG	4000~400	4000~400	RidgeR	RFR	0.64	0.90	1.55	3.20	0.61	0.80	1.46	2.20
C14:1	SNV	MSC	3017~2823/1805~1734	3017~2823/1805~1734	LassoR	PLSR	0.63	0.38	1.65	1.28	0.51	0.62	1.40	1.63
C16:1	SNV	MSC	3017~2823/1805~1734	3017~2823/1805~1734	LassoR	LassoR	0.54	0.38	1.47	1.27	0.62	0.55	1.64	1.50
C18:1n9c	SG	MSC	3017~2823/1805~1734	3017~2823/1805~1734	LassoR	LassoR	0.60	0.52	1.58	1.45	0.77	0.34	2.00	1.20
C20:1	SG	DER2	3017~2823/1805~1734	4000~400	PLSR	RFR	0.54	0.77	1.48	2.06	0.49	0.76	1.41	2.04
C22:1n9	DER2	DER1	4000~400	4000~400	RFR	RFR	0.51	0.53	1.43	1.45	0.45	0.65	1.36	1.67
C18:2n6c	MSC	SG	4000~400	4000~400	RFR	RidgeR	0.59	0.13	1.56	1.07	0.63	0.15	1.61	1.08
C18:3n3	SG	DER1	4000~400	3017~2823/1805~1734	RFR	RFR	0.60	0.17	1.59	1.09	0.70	0.27	1.82	1.13
C18:3n6	SG	DER2	3017~2823/1805~1734	3017~2823/1805~1734	RFR	RFR	0.18	0.84	1.08	2.47	0.14	0.76	1.04	2.00
C20:3n6	DER1	MSC	4000~400	3017~2823/1805~1734	RFR	PLSR	0.50	0.23	1.42	1.15	0.62	0.39	1.61	1.29
C20:4n6	SNV	SNV	4000~400	4000~400	RFR	PLSR	0.44	0.29	1.34	1.19	0.50	0.46	1.42	1.37
C20:5n3	DER2	DER1	4000~400	4000~400	RFR	RFR	0.33	0.83	1.23	2.41	0.43	0.91	1.29	3.06
LCFA	DER2	DER1	3017~2823/1805~1734	4000~400	RidgeR	RFR	0.68	0.41	1.78	1.31	0.83	0.42	2.29	1.32
MCFA	DER2	SNV	3017~2823/1805~1734	3017~2823/1805~1734	RFR	LassoR	0.64	0.23	1.67	1.14	0.75	0.28	2.00	1.18
MUFA	DER2	DER1	3017~2823/1805~1734	3017~2823/1805~1734	LassoR	LassoR	0.61	0.56	1.59	1.51	0.79	0.43	2.06	1.30
PUFA	DER2	SG	4000~400	3017~2823/1805~1734	RidgeR	RFR	0.71	0.16	1.80	1.08	0.71	0.16	1.75	1.07
SCFA	DER2	MSC	4000~400	3017~2823/1805~1734	RFR	LassoR	0.66	0.51	1.71	1.43	0.77	0.48	2.04	1.37
SFA	SG	SG	3017~2823/1805~1734	3017~2823/1805~1734	RFR	LassoR	0.66	0.32	1.73	1.21	0.76	0.25	2.01	1.16
UFA	DER2	MSC	3017~2823/1805~1734	3017~2823/1805~1734	LassoR	LassoR	0.62	0.42	1.62	1.31	0.82	0.48	2.15	1.38

Note: MIRS, DER1, DER2, MSC, SNV, SG, RFR, PLSR, LassoR, RidgeR,  $R^2$ , and RPD indicate mid-infrared spectrum, first-order derivative, second-order derivative, multiple scattering correction, standard normal transform, Savitzky-Golay convolution smoothing, random forest regression, partial least square regression, least absolute shrinkage and selection operator regression, ridge regression, determination coefficient, and residual predictive deviation, respectively.



**Figure 1.** MIRS after DER1, DER2, and SG pre-processing algorithms. Note: MIRS, DER1, DER2, and SG indicate mid-infrared spectrum, first-order derivative, second-order derivative, and Savitzky–Golay convolution smoothing, respectively.

### 3. Materials and Methods

#### 3.1. Milk Samples and Fatty Acids

Milk samples were collected from 336 Holstein cows on a farm in Shandong Province, China, including one small tube (30 mL) and one large tube (50 mL) from each cow. After sampling, all tubes were immediately stored in liquid nitrogen ( $-196\text{ }^{\circ}\text{C}$ ) and delivered to our experimental lab for further analysis within 6 h. In this study, to maintain analysis consistency, none of the 672 collected milk samples received any preservative additions, and the milk samples in the 30 mL and 50 mL tubes were used to measure fatty acid contents and MIRS, respectively.

A total of 24 kinds of fatty acid contents, which included C8:0, C10:0, C11:0, C12:0, C13:0, C14:0, C14:1, C15:0, C16:0, C16:1, C17:0, C18:0, C18:1n9c, C18:2n6c, C20:0, C18:3n6, C18:3n3, C20:1, C22:0, C20:3n6, C20:4n6, C22:1n9, C20:5n3, and C24:0, were measured and quantified in each milk sample from the 30 mL tubes using the GC technique. Due to the limitations of GC technique, the apparent missing values were replaced by the averaged values of the whole fatty acids that had been quantified by the GC technique. The outliers of quality control for the fatty acids were defined by the mean reference values  $\pm$  two standard deviations. For each milk sample from the 50 mL tubes, 899 raw data points for MIRS values in the complete waveband range of  $4000\sim 400\text{ cm}^{-1}$  were obtained by Bentley spectrometers (Bentley Instruments Inc., Chaska, MN, United States), following the routine methodology (e.g., 30 min preheating and sufficient shaking before operation). Afterwards, additional raw MIRS values, as the measurement replicates, were also obtained using the same milk samples. Finally, two raw MIRS values were transformed by the Fourier method [24] for further pre-processing steps.

Here, the GC methodology for the quantification of fatty acid contents in our study was similar to those in other studies [4,25]. The outputs of the GC technique were generated by analyzing the methyl esters from the fat in the milk following ISO Standard 15884 (ISO–IDF (International Organization for Standardization–International Dairy Federation),

2002). Normally, the GC technique is used as the gold standard for fatty acid measurements because of its high accuracy, even for low contents [26,27], while the MIRS method is more rapid and less expensive [13,21].

According to the saturation conditions of hydrocarbon chains, fatty acids are classified as saturated fatty acids (SFAs), unsaturated fatty acids (UFAs), monounsaturated fatty acids (MUFAs), and polyunsaturated fatty acids (PUFAs) [5]. According to the carbon chain lengths, fatty acids are classified as short chain (4 to 10 carbons) fatty acids (SCFAs), medium chain (11 to 16 carbons) fatty acids (MCFAs), and long chain (more than 16 carbons) fatty acids (LCFAs). Consequently, 7 fatty acid groups for 24 kinds of the above fatty acids were obtained (Table 4).

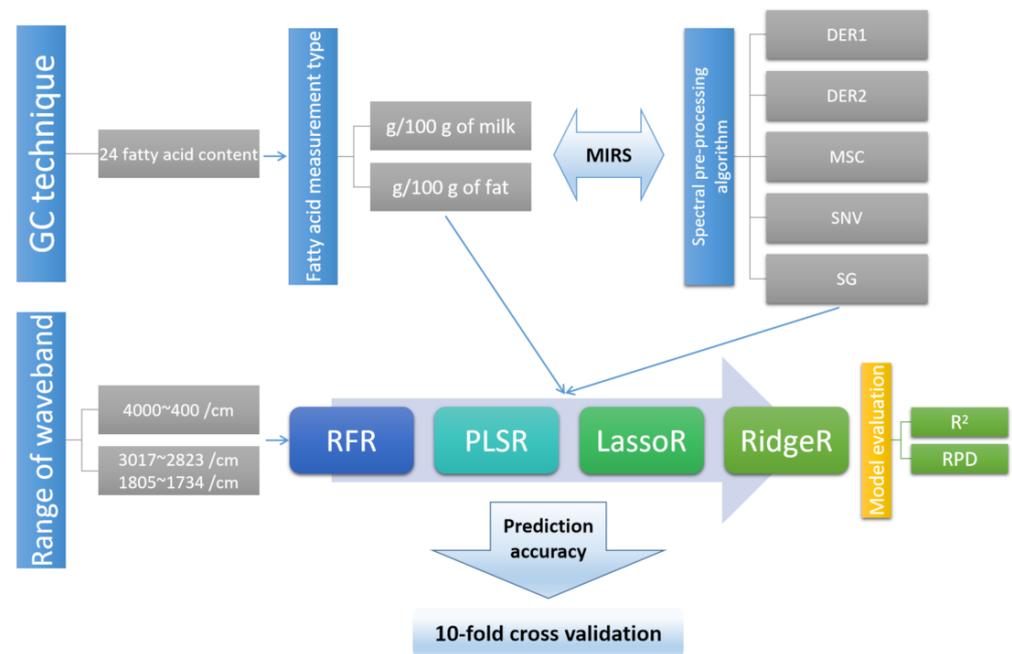
**Table 4.** Seven classified fatty acid groups, according to hydrocarbon chain saturation and carbon chain length.

Fatty Acid Group According to Hydrocarbon Chain Saturation		Fatty Acid Group According to Carbon Chain Length	
SFA	C8:0, C10:0, C11:0, C12:0, C13:0, C14:0, C15:0, C16:0, C17:0, C18:0, C20:0, C22:0, C24:0	SCFA	C8:0, C10:0
UFA	C14:1, C16:1, C18:1n9c, C18:2n6c, C18:3n6, C18:3n3, C20:1, C20:3n6, C20:4n6, C22:1n9, C20:5n3	MCFA	C11:0, C12:0, C13:0, C14:0, C15:0, C16:0, C16:1
MUFA	C14:1, C16:1, C18:1n9c, C20:1, C22:1n9	LCFA	C17:0, C18:0, C18:1n9c, C18:2n6c, C20:0, C18:3n6, C18:3n3, C20:1, C22:0, C20:3n6, C20:4n6, C22:1n9, C20:5n3, C24:0
PUFA	C18:2n6t, C18:2n6c, C18:3n6, C18:3n3, C20:3n6, C20:4n6, C22:2, C20:5n3		

Note: SFA, UFA, MUFA, PUFA, SCFA, MCFA, and LCFA indicate saturated fatty acid, unsaturated fatty acid, monounsaturated fatty acid, polyunsaturated fatty acid, short chain (4 to 10 carbons) fatty acid, medium chain (11 to 16 carbons) fatty acid, and long chain (more than 16 carbons) fatty acid, respectively.

### 3.2. Predictions of Milk Fatty Acid Contents Using MIRS Data

Each fatty acid content quantified by the GC technique was converted from g/100 g of milk (milk-basis) to g/100 g of fat (fat-basis) using the fat contents determined by MIRS. The final MIRS values (the averaged values of two transformed MIRS replicates using the same milk sample) were processed using five spectral pre-processing algorithms, i.e., first-order derivative (DER1), second-order derivative (DER2), multiple scattering correction (MSC), standard normal transform (SNV), and Savitzky–Goltsy convolution smoothing (SG). In order to compare the influence of each pre-processing algorithm, we used them individually to process the final MIRS values. Two types of fatty acid measurements (g/100 g of milk and g/100 g fat), with the five pre-processed spectra above and two ranges of wavebands ( $4000\sim 400\text{ cm}^{-1}$  and  $3017\sim 2823\text{ cm}^{-1}/1805\sim 1734\text{ cm}^{-1}$ ), were analyzed using four regression models, i.e., random forest regression (RFR), partial least square regression (PLSR), least absolute shrinkage and selection operator regression (LassoR), and ridge regression (RidgeR). The determination coefficient ( $R^2$ ) and residual predictive deviation (RPD) were used to evaluate the metrics of the four regression models. Prediction accuracy was assessed using a 10-fold cross validation procedure with the ratio of the training set and the test set as 3:1. The GC quantification technique, fatty acid measurements, spectral pre-processing algorithms, fatty acid prediction methods, and prediction accuracy assessments are summarized in Figure 2.



**Figure 2.** Summary of fatty acid prediction methods. Note: GC, MIRS, DER1, DER2, MSC, SNV, SG, RFR, PLSR, LassoR, RidgeR, R<sup>2</sup>, and RPD indicate gas chromatography, mid-infrared spectrum, first-order derivative, second-order derivative, multiple scattering correction, standard normal transform, Savitzky–Golay convolution smoothing, random forest regression, partial least square regression, least absolute shrinkage and selection operator regression, ridge regression, determination coefficient, and residual predictive deviation, respectively.

#### 4. Discussion

The concentrations of different milk fatty acids in our study (Table 1) seem slightly lower than those in other studies [5,28–30], which could be caused mainly by the differences in feed diet and milk-collection times of the farm, where they supplied their own total mix ration (TMR) three times per day, which is less than other similar Chinese Holstein cattle farms (four or five times per day). Compared to the results of Soyeurt et al. (2011) and Fleming et al. (2017), the variation coefficients ranged from 12.978% to 44.207% as fat-basis (g/100 g of fat), which were slightly lower, on average, than those in other studies. The higher variations of fatty acids as fat-basis (g/100 g of fat) in relation to those as milk-basis (g/100 g of milk) could be a tendency in which fatty acids exhibited high mean values and variation coefficients (Table 1).

Many previous studies have investigated the accuracy and applicability of prediction models based on R<sup>2</sup> values. Soyeurt et al. (2011) suggested that models with R<sup>2</sup> > 0.75 might be utilized for animal breeding. However, Zaalberg et al. (2021) used prediction models with R<sup>2</sup> > 0.6 for mineral elements in animal breeding [31]. Cecchinato et al. (2009) showed low R<sup>2</sup> values for curd characteristics predicted by MIRS, but they found high genetic correlations between the measured values and the predicted values [32]. In our study, 17 fatty acids (C8:0, C10:0, C12:0, C14:0, C18:0, C20:0, C22:0, C24:0, C18:1n9c, C20:1, C20:5n3, SFA, UFA, MUFA, SCFA, MCFA, and LCFA) showed RPD ≥ 2 and R<sup>2</sup> ≥ 0.75 (Table 2), which is consistent with the results of Soyeurt et al. (2006). This suggests that these 17 fatty acids can be accurately predicted using MIRS, and that this method has the potential for further fat trait selections in animal breeding. Furthermore, 6 fatty acids (C12:0, C20:0, C22:0, C20:5n3, UFA, and LCFA) with R<sup>2</sup> > 0.8, which were well predicted by MIRS, could also be used for breeding selections. For the grouped fatty acids, the R<sup>2</sup> values of the test set were greater than 0.7 (Table 3), which is consistent with the results of Soyeurt et al. (2006), Rutten et al. (2009), and Fleming et al. (2017). For both the training and the test sets, 6 individual fatty acids (C20:0, C22:0, C24:0, C20:1, C18:3n6, and C20:5n3) as fat-basis

(g/100 g of fat) showed  $R^2$  values greater than 0.7, whereas inconsistent results were found in other studies [4,5]. Fleming et al. (2017) obtained higher accuracy ( $R^2$ ) from fatty acids expressed on the milk-basis than on the fat-basis. Soyeurt et al. (2011) used the fatty acids predicted in milk for their prediction in fat and only achieved results better than those of the direct prediction in fat for C6:0, C12:0, C18:2 cis-9, cis-12, SFA, and SCFA. RPD is also used to measure the prediction effect and accuracy of models [33,34]. Three classifications of RPD are as follows: high prediction accuracy, which can be used for the quantitative prediction of substances when  $RPD \geq 2$ ; good prediction, which can be used for rough quantitative prediction or qualitative analysis when  $1.4 \leq RPD < 2$ ; and low prediction accuracy, which cannot be used for quantitative prediction when  $RPD < 1.4$ . Generally, a higher accuracy ( $R^2$  and RPD) can also be observed in the prediction of fatty acids by MIRS on the milk-basis ( $n = 22$ ) than on the fat-basis ( $n = 9$ ) (Tables 2 and 3), which is consistent with the results of other studies [4,5,21,35].

Different spectral pre-processing algorithms influence the prediction accuracy of fatty acids. Soyeurt et al. (2012) used MIRS to predict the lactoferrin content in bovine milk and obtained the highest prediction accuracy using PLSR based on DER1. Our study also found that derivatives (DER1 and DER2) and SG smoothing algorithms can be applied for most fatty acid predictions (Table 2). The derivative algorithm uses the absorbance values corresponding to each of two adjacent wave points to calculate their derivative values, where the spectrum is processed by the derivative. The wave points with large differences in absorbance reduce signal/noise interference; then, the corresponding value of the current wave point moves sequentially to retain the spectral information for stronger spectrum continuity (Figure 1).

## 5. Conclusions

In this study, different regression models led to varying prediction accuracy of fatty acid contents, while different pre-processing algorithms for the spectra also influenced prediction accuracy. It was revealed that a higher accuracy for most fatty acids can be achieved when derivative and SG pre-processing algorithms for RFR models were used. Therefore, after a series of evaluations in Chinese Holstein cows, these results suggest that the application of MIRS to predict the fatty acid contents of milk is feasible.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/molecules28020666/s1>, Supplementary File S1: All prediction accuracies (determination coefficient ( $R^2$ ) and residual predictive deviations (RPD)) after four regression model analysis (random forest regression (RFR), partial least square regression (PLSR), least absolute shrinkage and selection operator regression (LassoR), and ridge regression (RidgeR)) based on two types of fatty acid measurements (g/100 g of milk and g/100 g fat), two ranges of wavebands ( $4000\text{--}400\text{ cm}^{-1}$  and  $3017\text{--}2823\text{ cm}^{-1}/1805\text{--}1734\text{ cm}^{-1}$ ), and five spectral pre-processing algorithms (first-order derivative (DER1), second-order derivative (DER2), multiple scattering correction (MSC), standard normal transform (SNV), and Savitzky–Golay convolution smoothing (SG)).

**Author Contributions:** Conceptualization, J.L., X.W. and D.S.; methodology, X.Z. and Y.S.; formal analysis, X.Z. and Y.S.; resources, Y.Z., G.C., G.X., Y.L. and Y.G.; data curation, K.C., F.Z., K.W. and M.Z.; writing—original draft preparation, X.Z., X.W. and Y.S.; writing—review and editing, X.W.; supervision, J.L. and X.W.; funding acquisition, J.L. and D.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by Shandong Provincial Natural Science Foundation (ZR2021MC070), the National Key R&D Program of China (2021YFF1000701-06), Shandong Provincial Natural Science Foundation (ZR2020MC165), and the Earmarked Fund for CARS-36.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We thank the farm in Shandong Province, China, for providing the milk samples.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Sample Availability:** Samples of the fatty acid quantification and the mid-infrared spectrum (MIRS) value are available from the authors.

## References

1. Christie, W. The Composition and Structure of Milk Lipids. In *Developments in Dairy Chemistry—2*; Springer: Berlin/Heidelberg, Germany, 1983. [\[CrossRef\]](#)
2. Balaji, B.; Dehghan, M.; Mente, A.; Rangarajan, S.; Yusuf, S. Association of Dairy Consumption with Metabolic Syndrome, Hypertension and Diabetes in 147,812 Individuals from 21 Countries. *SSRN Electron. J.* **2019**. [\[CrossRef\]](#)
3. Haug, A.; Stmark, A.T.; Harstad, O.M. Bovine milk in human nutrition—A review. *Lipids Health Dis.* **2007**, *6*, 25. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Soyeurt, H.; Dehareng, F.; Gengler, N.; McParland, S.; Wall, E.; Berry, D.P.; Coffey, M.; Dardenne, P. Mid-infrared prediction of bovine milk fatty acids across multiple breeds, production systems, and countries. *J. Dairy Sci.* **2011**, *94*, 1657–1667. [\[CrossRef\]](#)
5. Fleming, A.; Schenkel, F.S.; Chen, J.; Malchiodi, F.; Bonfatti, V.; Ali, R.A.; Mallard, B.; Corredig, M.; Miglior, F. Prediction of milk fatty acid content with mid-infrared spectroscopy in Canadian dairy cattle using differently distributed model development sets. *J. Dairy Sci.* **2017**, *100*, 5073–5081. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Jensen, R.G. The composition of bovine milk lipids: January 1995 to December 2000. *J. Dairy Sci.* **2002**, *85*, 295–350. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Stoop, W.M.; Bovenhuis, H.; Heck, J.M.L.; van Arendonk, J.A.M. Effect of lactation stage and energy status on milk fat composition of Holstein-Friesian cows. *J. Dairy Sci.* **2009**, *92*, 1469–1478. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Wang, F.; Chen, M.; Luo, R.; Huang, G.; Wu, X.; Zheng, N.; Zhang, Y.; Wang, J. Fatty acid profiles of milk from Holstein cows, Jersey cows, buffalos, yaks, humans, goats, camels, and donkeys based on gas chromatography–mass spectrometry. *J. Dairy Sci.* **2021**, *105*, 1687–1700. [\[CrossRef\]](#)
9. Grelet, C.; Dardenne, P.; Soyeurt, H.; Fernandez, J.A.; Vanlierde, A.; Stevens, F.; Gengler, N.; Dehareng, F. Large-scale phenotyping in dairy sector using milk MIR spectra: Key factors affecting the quality of predictions. *Methods* **2021**, *186*, 97–111. [\[CrossRef\]](#)
10. Rovere, G.; de Los Campos, G.; Lock, A.; Worden, L.; Vazquez, A.I.; Lee, K.; Tempelman, R.J. Prediction of fatty acid composition using milk spectral data and its associations with various mid-infrared spectral regions in Michigan Holsteins. *J. Dairy Sci.* **2021**, *104*, 11242–11258. [\[CrossRef\]](#)
11. Liu, X.M.; Zhang, Y.; Zhou, Y.; Li, G.H.; Feng, X.S. Progress in Pretreatment and Analysis of Fatty Acids in Foods: An Update since 2012. *Sep. Purif. Rev.* **2019**, *50*, 203–222. [\[CrossRef\]](#)
12. Waktola, H.D.; Zeng, A.X.; Chin, S.-T.; Marriott, P.J. Advanced gas chromatography and mass spectrometry technologies for fatty acids and triacylglycerols analysis. *Trends Anal. Chem.* **2020**, *129*, 115957. [\[CrossRef\]](#)
13. Samková, E.; Špička, J.; Hanuš, O.; Roubal, P.; Pecová, L.; Hasoňová, L.; Smetana, P.; Klimešová, M.; Čítek, J. Comparison of Fatty Acid Proportions Determined by Mid-Infrared Spectroscopy and Gas Chromatography in Bulk and Individual Milk Samples. *Animals* **2020**, *10*, 1095. [\[CrossRef\]](#)
14. Lopez-Villalobos, N.; Spelman, R.; Melis, J.; Davis, S.; Berry, S.; Lehnert, K.; Sneddon, N.; Holroyd, S.; MacGibbon, A.; Snell, R. Genetic correlations of milk fatty acid contents predicted from milk mid-infrared spectra in New Zealand dairy cattle. *J. Dairy Sci.* **2020**, *103*, 7238–7248. [\[CrossRef\]](#)
15. Du, C.; Nan, L.; Yan, L.; Bu, Q.; Ren, X.; Zhang, Z.; Sabek, A.; Zhang, S. Genetic Analysis of Milk Production Traits and Mid-Infrared Spectra in Chinese Holstein Population. *Animals* **2020**, *10*, 139. [\[CrossRef\]](#)
16. Ferrand-Calmels, M.; Palhière, I.; Brochard, M.; Leray, O.; Astruc, J.; Aurel, M.; Barbey, S.; Bouvier, F.; Brunschwig, P.; Caillat, H.; et al. Prediction of fatty acid profiles in cow, ewe, and goat milk by mid-infrared spectrometry. *J. Dairy Sci.* **2014**, *97*, 17–35. [\[CrossRef\]](#)
17. Caredda, M.; Addis, M.; Ibba, I.; Leardi, R.; Scintu, M.F.; Piredda, G.; Sanna, G. Prediction of fatty acid content in sheep milk by Mid-Infrared spectrometry with a selection of wavelengths by Genetic Algorithms. *LWT Food Sci. Technol.* **2016**, *65*, 503–510. [\[CrossRef\]](#)
18. Caredda, M.; Addis, M.; Ibba, I.; Leardi, R.; Scintu, M.F.; Piredda, G.; Sanna, G. Building of prediction models by using Mid-Infrared spectroscopy and fatty acid profile to discriminate the geographical origin of sheep milk. *LWT Food Sci. Technol.* **2017**, *75*, 131–136. [\[CrossRef\]](#)
19. Nakayama, J.Y.; Ho, J.; Cartwright, E.; Simpson, R.; Hertzberg, V.S. Predictors of progression through the cascade of care to a cure for hepatitis C patients using decision trees and random forests. *Comput. Biol. Med.* **2021**, *134*, 104461. [\[CrossRef\]](#)
20. Wang, F.; Wang, Y.; Zhang, K.; Hu, M.; Weng, Q.; Zhang, H. Spatial heterogeneity modeling of water quality based on random forest regression and model interpretation. *Environ. Res.* **2021**, *202*, 111660. [\[CrossRef\]](#)
21. Soyeurt, H.; Dardenne, P.; Dehareng, F.; Lognay, G.; Veselko, D.; Marlier, M.; Bertozzi, C.; Mayeres, P.; Gengler, N. Estimating Fatty Acid Content in Cow Milk Using Mid-Infrared Spectrometry. *J. Dairy Sci.* **2006**, *89*, 3690–3695. [\[CrossRef\]](#)
22. Soyeurt, H.; Bastin, C.; Colinet, F.G.; Arnould, V.M.R.; Berry, D.P.; Wall, E.; Dehareng, F.; Nguyen, H.N.; Dardenne, P.; Schefers, J.; et al. Mid-infrared prediction of lactoferrin content in bovine milk: Potential indicator of mastitis. *Animal* **2012**, *6*, 1830–1838. [\[CrossRef\]](#)

23. Ho, P.N.; Marett, L.C.; Wales, W.J.; Axford, M.; Oakes, E.M.; Pryce, J.E. Predicting milk fatty acids and energy balance of dairy cows in Australia using milk mid-infrared spectroscopy. *Anim. Prod. Sci.* **2019**, *60*, 164. [[CrossRef](#)]
24. Tiplady, K.M.; Lopdell, T.J.; Littlejohn, M.D.; Garrick, D.J. The evolving role of Fourier-transform mid-infrared spectroscopy in genetic improvement of dairy cattle. *J. Anim. Sci. Biotechnol.* **2020**, *11*, 39. [[CrossRef](#)]
25. Soyeurt, H.; Dehareng, F.; Mayeres, P.; Bertozzi, C.; Gengler, N. Variation of delta9-desaturase activity in dairy cattle. *J. Dairy Sci.* **2008**, *91*, 3211–3224. [[CrossRef](#)]
26. Christie, W. Gas chromatography mass spectrometry methods for structural analysis of fatty acids. *Lipids* **1998**, *33*, 343–353. [[CrossRef](#)]
27. Maurice-van Eijndhoven, M.H.T.; Soyeurt, H.; Dehareng, F.; Calus, M. Validation of fatty acid predictions in milk using mid-infrared spectrometry across cattle breeds. *Animal* **2013**, *7*, 348–354. [[CrossRef](#)]
28. Coppa, M.; Ferlay, A.; Leroux, C.; Jestin, M.; Chilliard, Y.; Martin, B.; Andueza, D. Prediction of milk fatty acid composition by near infrared reflectance spectroscopy. *Int. Dairy J.* **2010**, *20*, 182–189. [[CrossRef](#)]
29. De Marchi, M.; Penasa, M.; Cecchinato, A.; Mele, M.; Secchiari, P.; Bittante, G. Effectiveness of mid-infrared spectroscopy to predict fatty acid composition of Brown Swiss bovine milk. *Animal* **2011**, *5*, 1653–1658. [[CrossRef](#)]
30. De Marchi, M.; Toffanin, V.; Cassandro, M.; Penasa, M. Invited review: Mid-infrared spectroscopy as phenotyping tool for milk traits. *J. Dairy Sci.* **2014**, *97*, 1171–1186. [[CrossRef](#)]
31. Zaalberg, R.M.; Poulsen, N.A.; Bovenhuis, H.; Sehested, J.; Larsen, L.B.; Buitenhuis, A.J. Genetic analysis on infrared-predicted milk minerals for Danish dairy cattle. *J. Dairy Sci.* **2021**, *104*, 8947–8958. [[CrossRef](#)]
32. Cecchinato, A.; Marchi, M.D.; Gallo, L.; Bittante, G.; Carnier, P. Mid-infrared spectroscopy predictions as indicator traits in breeding programs for enhanced coagulation properties of milk. *J. Dairy Sci.* **2009**, *92*, 5304–5313. [[CrossRef](#)] [[PubMed](#)]
33. Canaza-Cayo, A.W.; Alomar, D.; Quispe, E. Prediction of alpaca fibre quality by near-infrared reflectance spectroscopy. *Anim. Int. J. Anim. Biosci.* **2013**, *7*, 1219–1225. [[CrossRef](#)] [[PubMed](#)]
34. Pralle, R.S.; Weigel, K.W.; White, H.M. Predicting blood  $\beta$ -hydroxybutyrate using milk Fourier transform infrared spectrum, milk composition, and producer-reported variables with multiple linear regression, partial least squares regression, and artificial neural network. *J. Dairy Sci.* **2018**, *101*, 4378–4387. [[CrossRef](#)] [[PubMed](#)]
35. Rutten, M.J.M.; Bovenhuis, H.; Hettinga, K.A.; van Valenberg, H.J.F.; van Arendonk, J.A.M. Predicting bovine milk fat composition using infrared spectroscopy based on milk samples collected in winter and summer. *J. Dairy Sci.* **2009**, *92*, 6202–6209. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.