Supplemental Table 1. Indicators of all developed models for buffalo raw milk.

Algorithm	Pre-processed of MIR[1]	Dataset	Indicators					
			Accuracy	Sensitivity	Specificity	AUC[2]	F1 score	MCC
Partial least squares discriminant analysis (PCs: 20)[3]	Original	Calibration	0.95	0.94	1.00	0.99	0.98	0.82
		Validation	0.85	0.82	1.00	0.96	0.97	0.67
	Diff	Calibration	0.93	0.92	1.00	0.99	0.98	0.79
		Validation	0.80	0.76	1.00	0.94	0.94	0.60
	SNV	Calibration	0.94	0.93	1.00	0.99	0.97	0.81
		Validation	0.88	0.85	1.00	0.94	0.95	0.71
	MSC	Calibration	0.95	0.94	1.00	0.99	0.99	0.84
		Validation	0.83	0.79	1.00	0.95	0.95	0.63
	SG(13,4)	Calibration	0.96	0.96	1.00	1.00	0.99	0.87
		Validation	0.85	0.82	1.00	0.97	0.98	0.67
Support vector machine[4]	**Original**	**Calibration**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
		Validation	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
	Diff	Calibration	0.93	0.98	0.63	0.95	0.96	0.68
		Validation	0.86	0.92	0.50	0.93	0.92	0.42
	SNV	Calibration	1.00	1.00	1.00	1.00	1.00	1.00
		Validation	0.93	0.97	0.67	0.88	0.96	0.69
	MSC	Calibration	0.98	0.99	0.88	0.98	0.99	0.90
		Validation	0.90	0.94	0.67	0.96	0.94	0.61
	SG(17,2)	Calibration	0.99	1.00	0.96	1.00	1.00	0.98
		Validation	0.98	0.97	1.00	1.00	0.99	0.91
Random forest	Original	Calibration	1.00	1.00	1.00	1.00	1.00	1.00
		Validation	0.83	0.86	0.67	0.88	0.90	0.45
	Diff	Calibration	1.00	1.00	1.00	1.00	1.00	1.00
		Validation	0.86	0.92	0.50	0.87	0.92	0.42
	SNV	Calibration	1.00	1.00	1.00	1.00	1.00	1.00
		Validation	0.83	0.92	0.33	0.83	0.90	0.27
	MSC	Calibration	1.00	1.00	1.00	1.00	1.00	1.00
		Validation	0.81	0.86	0.50	0.88	0.89	0.32
	SG(17,2)	Calibration	1.00	1.00	1.00	1.00	1.00	1.00
		Validation	0.83	0.89	0.50	0.90	0.90	0.37

[1]: MIR: mid-infrared spectra; Diff: first-order difference; SNV: standardized normal variation; MSC: multiplicative scatter correction; SG: Savitzky-Golag (window length, poly order)

[2]: AUC: area under curve

[3]: PCs: number of principal components used by partial least squares discriminant analysis

[4]: Bold indicates the optimal model for this type of milk

Supplemental Table 2. Indicators of all developed models for bovine raw milk.

Algorithm	Pre-processed of MIR[1]	Dataset	Indicators					
			Accuracy	Sensitivity	Specificity	AUC[2]	F1 score	MCC
Partial least squares discriminant analysis (PCs: 20)[3,4]	Original	Calibration	1.00	1.00	1.00	1.00	1.00	1.00
		Validation	0.96	0.94	1.00	0.99	0.99	0.90
	Diff	Calibration	1.00	1.00	1.00	1.00	1.00	1.00
		Validation	0.92	1.00	0.67	0.98	0.98	0.77
	SNV	Calibration	1.00	1.00	1.00	1.00	1.00	1.00
		Validation	0.96	0.94	1.00	0.99	0.99	0.90
	MSC	Calibration	1.00	1.00	1.00	1.00	1.00	1.00
		Validation	0.96	0.94	1.00	0.99	0.99	0.90
	SG(13,7)	Calibration	1.00	1.00	1.00	1.00	1.00	1.00
		Validation	0.92	0.94	0.83	0.98	0.98	0.78
Support vector machine	Original	Calibration	0.99	0.99	1.00	1.00	0.99	0.97
		Validation	0.96	1.00	0.83	1.00	0.97	0.89
	Diff	**Calibration**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
		Validation	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
	SNV	Calibration	1.00	1.00	1.00	1.00	1.00	1.00
		Validation	0.79	0.89	0.50	0.84	0.86	0.41
	MSC	Calibration	0.98	0.99	0.96	1.00	0.99	0.94
		Validation	0.83	0.89	0.67	0.90	0.89	0.56
	SG(17,2)	Calibration	1.00	1.00	1.00	1.00	1.00	1.00
		Validation	0.88	1.00	0.50	0.84	0.92	0.65
Random forest	Original	Calibration	1.00	1.00	1.00	1.00	1.00	1.00
		Validation	0.79	0.94	0.33	0.69	0.87	0.36
	Diff	Calibration	1.00	1.00	1.00	1.00	1.00	1.00
		Validation	0.83	1.00	0.33	0.80	0.90	0.52
	SNV	Calibration	1.00	1.00	1.00	1.00	1.00	1.00
		Validation	0.71	0.94	0.00	0.28	0.83	-0.12
	MSC	Calibration	1.00	1.00	1.00	1.00	1.00	1.00
		Validation	0.58	0.78	0.00	0.31	0.74	-0.26
	SG(17,2)	Calibration	1.00	1.00	1.00	1.00	1.00	1.00
		Validation	0.79	0.94	0.33	0.82	0.87	0.36

[1]: MIR: mid-infrared spectra; Diff: first-order difference; SNV: standardized normal variation; MSC: multiplicative scatter correction; SG: Savitzky-Golag (window length, poly order)

[2]: AUC: area under curve

[3]: PCs: number of principal components used by partial least squares discriminant analysis

[4]: Bold indicates the optimal model for this type of milk

Supplemental Table 3. Indicators of all developed models for bovine pasteurized milk.

Algorithm	Pre-processed of MIR[1]	PCs[2]	Dataset	Indicators		
				Accuracy	kappa	F1 score
Partial least squares discriminant analysis	Original	5	Calibration	0.15	-0.10	0.10
			Validation	0.18	-0.10	0.11
	Diff	3	Calibration	0.14	-0.09	0.09
			Validation	0.16	-0.11	0.09
	SNV	2	Calibration	0.29	0.04	0.20
			Validation	0.32	0.04	0.18
	MSC	2	Calibration	0.21	-0.04	0.15
			Validation	0.24	-0.02	0.15
	SG(27,2)	2	Calibration	0.20	-0.03	0.11
			Validation	0.26	0.01	0.13
Support vector machine	Original	NA[3]	Calibration	0.96	0.95	0.96
			Validation	0.75	0.63	0.75
	Diff		Calibration	0.86	0.79	0.86
			Validation	0.72	0.58	0.71
	SNV		Calibration	0.91	0.87	0.91
			Validation	0.59	0.39	0.58
	MSC		Calibration	0.84	0.76	0.84
			Validation	0.59	0.39	0.58
	SG(27,3)		Calibration	0.82	0.73	0.82
			Validation	0.77	0.65	0.75
Random forest[4]	Original	NA	Calibration	1.00	1.00	1.00
			Validation	0.84	0.77	0.84
	Diff		Calibration	1.00	1.00	1.00
			Validation	0.86	0.79	0.86
	SNV		Calibration	1.00	1.00	1.00
			Validation	0.39	0.08	0.39
	MSC		Calibration	1.00	1.00	1.00
			Validation	0.34	0.02	0.34
	SG(27,3)		**Calibration**	**1.00**	**1.00**	**1.00**
			Validation	**0.94**	**0.91**	**0.94**

[1]: MIR: mid-infrared spectra; Diff: first-order difference; SNV: standardized normal variation; MSC: multiplicative scatter correction; SG: Savitzky-Golag (window length, poly order)

[2]: PCs: number of principal components used by partial least squares discriminant analysis

[3]: NA: not applicable

[4]: Bold indicates the optimal model for this type of milk

Supplemental Table 4. Indicators of all developed models for bovine ultra-high temperature sterilized (UHT) milk.

Algorithm	Pre-processed of MIR[1]	PCs[2]	Dataset	Indicators		
				Accuracy	kappa	F1 score
Partial least squares discriminant analysis	Original	2	Calibration	0.21	-0.10	0.10
			Validation	0.24	-0.05	0.15
	Diff	2	Calibration	0.27	0.05	0.15
			Validation	0.30	0.08	0.16
	SNV	2	Calibration	0.15	-0.04	0.10
			Validation	0.19	-0.01	0.12
	MSC	5	Calibration	0.18	-0.03	0.13
			Validation	0.22	0.01	0.16
	SG(21,2)	2	Calibration	0.30	0.07	0.17
			Validation	0.31	0.09	0.17
Support vector machine	Original	NA[3]	Calibration	0.99	0.99	1.00
			Validation	0.92	0.89	0.93
	Diff		Calibration	1.00	1.00	1.00
			Validation	0.90	0.86	0.91
	SNV		Calibration	0.97	0.95	0.97
			Validation	0.87	0.82	0.87
	MSC		Calibration	0.96	0.94	0.96
			Validation	0.87	0.82	0.87
	SG(21,2)		Calibration	1.00	1.00	1.00
			Validation	0.88	0.83	0.88
Random forest[4]	Original	NA	Calibration	1.00	1.00	1.00
			Validation	0.88	0.84	0.88
	Diff		Calibration	1.00	1.00	1.00
			Validation	0.92	0.89	0.93
	SNV		Calibration	1.00	1.00	1.00
			Validation	0.51	0.31	0.50
	MSC		Calibration	1.00	1.00	1.00
			Validation	0.45	0.21	0.35
	SG(21,2)		**Calibration**	**1.00**	**1.00**	**1.00**
			Validation	**0.95**	**0.92**	**0.95**

[1]: MIR: mid-infrared spectra; Diff: first-order difference; SNV: standardized normal variation; MSC: multiplicative scatter correction; SG: Savitzky-Golag (window length, poly order)

[2]: PCs: number of principal components used by partial least squares discriminant analysis

[3]: NA: not applicable

[4]: Bold indicates the optimal model for this type of milk