Accuracy and biases in predicting the chemical and physical traits of many types of cheeses using different visible and near-infrared spectroscopic techniques and spectrum intervals

Giorgia Stocco,[1,2]***** **Claudio Cipolat-Gotet,**[2] **Alessandro Ferragina,**[1] **Paolo Berzaghi,**[3] **and Giovanni Bittante**[1]
[1]Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE), University of Padova, Viale dell'Università 16, 35020 Legnaro, Italy
[2]Department of Veterinary Science, University of Parma, Via del Taglio 10, 43126 Parma, Italy
[3]Department of Animal Medicine, Production and Health (MAPS), University of Padova, Viale dell'Università 16, 35020 Legnaro, Italy

ABSTRACT

Near-infrared spectroscopy (NIRS) has been widely used to determine various composition traits of many dairy products in the industry. In the last few years, near-infrared (NIR) instruments have become more and more accessible, and now, portable devices can be easily used in the field, allowing the direct measurement of important quality traits. However, the comparison of the predictive performances of different NIR instruments is not simple, and the literature is lacking. These instruments may use different wavelength intervals and calibration procedures, making it difficult to establish whether differences are due to the spectral interval, the chemometric approach, or the instrument's technology. Hence, the aims of this study were (1) to evaluate the prediction accuracy of chemical contents (5 traits), pH, texture (2 traits), and color (5 traits) of 37 categories of cheese; (2) to compare 3 instruments [2 benchtop, working in reflectance (R) and transmittance (T) mode (NIRS-R and NIRS-T, respectively) and 1 portable device (VisNIRS-R)], using their entire spectral ranges (1100–2498, 850–1048, and 350–1830 nm, respectively, for NIRS-R, NIRS-T and VisNIRS-R); (3) to examine different wavelength intervals of the spectrum within instrument, comparing also the common intervals among the 3 instruments; and (4) to determine the presence of bias in predicted traits for specific cheese categories. A Bayesian approach was used to develop 8 calibration models for each of 13 traits. This study confirmed that NIR spectroscopy can be used to predict the chemical composition of a large number of different cheeses, whereas pH and texture traits were poorly predicted. Color showed variable predictability, according to the trait considered, the instrument used, and, within instrument, according to the wavelength intervals. The predictive performance of the VisNIRS-R portable device was generally better than the 2 laboratory NIRS instruments, whether with the entire spectrum or selected intervals. The VisNIRS-R was found suitable for analyzing chemical composition in real time, without the need for sample uptake and processing. Our results also indicated that instrument technology is much more important than the NIR spectral range for accurate prediction equations, but the visible range is useful when predicting color traits, other than lightness. Specifically for certain categories (i.e., caprine, moldy, and fresh cheeses), dedicated calibrations seem to be needed to obtain unbiased and more accurate results.

Key words: cheese quality, water-soluble protein, texture trait, Bayesian calibration, chemometric

INTRODUCTION

Infrared (**IR**) spectroscopy is an appealing technique for assessing food quality in numerous sectors, including dairy production (Karoui et al., 2010); the instruments can now be easily used in the field and are therefore increasingly less confined to the laboratory (Cattaneo and Holroyd, 2013). Both near-infrared (**NIR**) and mid-infrared (**MIR**) spectroscopy (**NIRS** and **MIRS**, respectively) have been used to monitor raw materials, processing, and final products (Woodcock et al., 2008; Berzaghi and Riovanto, 2009; Karoui et al., 2010).

The use of IR spectroscopy to predict cheese quality has been widely reviewed (Woodcock et al., 2008; Karoui et al., 2010). Most of these studies focused on predictions of fat, protein, and moisture contents (McQueen et al., 1995; Rodriguez-Otero et al., 1995; Lee et al., 1997). Many attempts were also made to predict some of the sensory and rheological properties of cheese, such as flavor and texture characteristics (Sørensen and Jepsen, 1998; Downey et al., 2005; Fagan et al., 2007a,b). Other research has addressed the

characterization of cheese by predicting several other chemical and physical traits, including water-soluble nitrogen (**WSN**), NaCl, vitamins, minerals, pH, and color (Lucas et al., 2008; González-Martín et al., 2011).

Much of the work concerning the potential of applying IR spectroscopy to cheese quality has been carried out on specific traits of particular cheese types using laboratory benchtop instruments that require sample preparation, such as grinding or removal of water (Pierce and Wehling, 1994; Downey et al., 2005; Karoui et al., 2010). Also, among results obtained on a specific cheese, it is not easy to compare the performances of different instruments. This is because they may use different wavelength intervals and calibration procedures, making it difficult to establish whether differences are due to the spectral interval, the chemometrics approach, or the instrument's technology.

Fast and more affordable NIR portable instruments have been now developed for their use at the field level (i.e., directly on the cheese surface). Moreover, new statistical approaches for calibration procedures have been devised for other food categories. To date, important advances have been made in developing Bayesian models for high-dimensional regressions, and many of these methods are adopted for regression on high-dimensional genotypes (i.e., genomic selection; Meuwissen et al., 2001). Models such as Bayesian generalized linear regression (**BGLR**) approaches, widely applied to animal and plant breeding, have been proved to have greater predictive power than dimension-reduction methods (i.e., partial least squares, **PLS**; de los Campos et al., 2013; Ferragina et al., 2015). This underlines the need to develop more efficient chemometric methods to analyze spectral data.

Hence, the objectives of our study were (1) to evaluate the predictability of several chemical and physical traits of a large variety of cheeses, by using a Bayesian chemometric procedure; (2) to compare 3 different instruments, 2 benchtop and 1 portable, using their entire spectral ranges; (3) to compare different spectral intervals, in the range of visible and NIR wavelengths, common to different instruments; and (4) to discover whether some of the quality traits of specific cheese categories are affected by biases.

MATERIALS AND METHODS

Origin and Collection of Cheese Samples

A total of 1,050 different cheeses from 104 cheese factories were sampled at 3 consecutive editions of Caseus Veneti, an annual cheese exhibition and competition (http://www.caseusveneti.it/) sponsored by the Veneto Regional Government (northeastern Italy) and

organized by the A.Pro.La.V. (Veneto Dairy Producers Association) with the aim of assessing and promoting the region's cheeses and producers. The Veneto dairy sector has a long tradition and produces a wide variety of cheeses. Because of their wide variability, the cheeses in competition are divided into 37 categories (Table 1), which can be classified into 3 main groups: PDO cheeses (16 categories, 321 cheeses), traditional cheeses (4 categories, 136 cheeses), and other cheeses (17 categories, 593 cheeses). The first group comprised only cheeses with protected designation of origin (**PDO**) status according to official European Union regulations (CEE 2081/92). Veneto traditionally produces 7 PDO cheeses, 2 of which comprised a single category each (Casatella Trevigiana PDO and Piave PDO cheeses). The other 5 were divided into more than one category, according to milk composition, manufacturing process, or length of ripening: Grana Padano PDO (2 categories), Provolone PDO (2 categories), Montasio PDO (3 categories), Monte Veronese PDO (3 categories), and Asiago PDO (4 categories).

The second group (Table 1) comprised cheeses recognized by the Veneto Regional Government as traditional cheeses, which were divided into 4 categories: Morlacco del Grappa cheese (produced on the Monte Grappa Massif), Malga cheeses (produced exclusively during the summer transhumance of cows to high Alpine pastures; 2 categories, according to length of ripening), and the Formaggio affinato nelle vinacce cheese (part of the ripening period is in must or wine).

The group of other cheeses included those that are produced without any officially recognized production regulations. These cheeses were divided into 17 categories (Table 1) according to the type of milk (i.e., goat cheese types), the addition of ingredients to the vat (i.e., pepper or chili cheeses category), curd handling (i.e., stretched cheese types), length of ripening (i.e., fresh soft cheeses, semi-hard cheeses, hard cheeses categories), cheese manufacture techniques allowing specific mold development (i.e., blue cheeses, moldy rind soft cheeses, washed rind soft cheeses categories).

At each edition of the Caseus Veneti contest, one cheese sample was collected immediately after cutting the cheese wheel, from all cheeses presented in the exhibition. One sample represented a different cheese (or group of cheeses of the same batch in the case of very small cheeses) type produced by a cheese factory participating in a given cheese category. Thus, the number of cheeses sampled for each category represents the number of cheese factories producing that type of cheese. The cheese samples were cooled to 4°C and transported to the Milk Quality Laboratory of the Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE) of the University of

Padova (Legnaro, Italy). All the samples were analyzed within 3 d of sampling, starting from categories that included fresh cheeses.

Chemical Analyses

The chemical composition (moisture, protein, lipids, ash, and WSN) was analyzed on a subsample of 197 selected cheeses from the third edition of Caseus Veneti and included in the database (Table 1). The cheeses were randomly selected from each category in a number varying according to the number and spectra variability of cheeses analyzed but always including the competition winners of each category. The chemical components were determined according to official analytical methods (AOAC, 1990): method 926.08 for moisture (vacuum oven method, drying at 102°C), method 935.42 for ash (gravimetric method, ashing at 550°C), ISO 14156:2001 for lipids (extraction method; ISO, 2001), and method 920.123 for CP and WSN (Kjeldahl method). Only these analyses were carried out on defrosted cheese samples stored at −20°C. The pH was measured in triplicate on all fresh cheeses using a portable pH meter (Crison Basic 25; Crison Instruments SA, Barcelona, Spain), and the 3 acquisitions were averaged before statistical analysis.

Texture Analysis

Texture traits of all the cheeses sampled in the second and third year of Caseus Veneti, excluding those fresh, soft cheeses, were determined using a texture analyzer

Table 1. Number of individual cheeses used for calibration of chemical composition, pH, texture, and color traits according to cheese category

Cheese category	Chemical composition	pH	Texture	Color
Total number of cheeses analyzed	197	1,042	602	1,046
PDO[1] cheeses	72	319	198	320
Asiago pressato (20 d–2 mo)	5	45	34	45
Asiago d'allevo mezzano (4–6 mo)	5	28	20	28
Asiago d'allevo vecchio (10–15 mo)	5	19	12	19
Asiago d'allevo stravecchio (>15 mo)	5	23	15	23
Casatella Trevigiana (very fresh, no crust)	5	27	4	26
Grana Padano (12–20 mo)	5	17	11	17
Grana Padano (>20 mo)	5	17	12	18
Montasio fresco (2–5 mo)	5	29	19	29
Montasio mezzano (5–10 mo)	5	23	15	23
Montasio stagionato (>10 mo)	5	20	12	20
Monte Veronese latte intero (25/45 d)	5	21	13	21
Monte Veronese d'allevo mezzano (3–6 mo)	5	18	10	18
Monte Veronese d'allevo vecchio (>12 mo)	4	13	7	13
Piave	2	6	4	6
Provolone Valpadana dolce	3	7	5	7
Provolone Valpadana piccante	3	6	5	7
Traditional cheeses	33	136	88	136
Morlacco del Grappa	7	30	21	30
Malga fresco (from summer Alpine pasture, 2–4 mo)	13	42	27	42
Malga vecchio (from summer Alpine pastures, >12 mo)	6	19	12	19
Formaggio affinato nelle vinacce (Marc must or wine refined)	7	45	28	45
Other cheeses	92	587	316	590
Freschissimi, pasta molle senza crosta (fresh soft cheeses, no crust)	7	55	8	57
Caciotta, pasta molle con crosta (soft fresh cheese with rind), <1 kg	4	55	36	55
Latteria, pasta molle con crosta (soft fresh cheeses with rind), >1 kg	7	76	49	76
Pasta molle con crosta fiorita (moldy rind soft cheeses)	4	25	12	25
Pasta molle con crosta lavata (washed rind soft cheeses)	5	18	9	18
Pasta semidura, 3–6 mesi (semi-hard cheese, 3–6 mo)	8	91	54	91
Pasta dura, oltre 10 mesi (hard cheese, >10 mo)	7	42	30	42
Mozzarella and Mozzarella TSG[2] cheese	6	20	8	20
Pasta filata molle (stretched-soft cheese)	3	13	1	13
Pasta filata dura (stretched-hard cheese)	3	11	6	11
Formaggi con pepe peperoncino (pepper or chili cheeses)	5	40	26	40
Formaggi con erbe fieno e spezie (herbs, hay, spices cheeses)	5	43	30	43
Formaggi affumicati (smoked cured cheeses)	5	22	14	22
Formaggi erborinati (blue cheeses)	6	19	9	19
Formaggi di capra, coag. acida (goat cheese, sour coagulation)	7	19	1	19
Formaggi di capra, coag. presamica (goat cheese, rennet coagulation)	10	34	19	35
Formaggi con birra o aceto bals. (beer, balsamic vinegar cheeses)	—	4	4	4

[1]PDO = protected designation of origin.

[2]TSG = traditional specialty guaranteed.

(XT2i, Stable Micro Systems Ltd., Godalming, Surrey, UK) with a Warner–Bratzler shear device (50-N load cell; 2 mm/s crosshead speed). For each cheese, 2 cylinder-shaped core samples were taken (1 cm^2 cross sectional area; 3 cm long), and their textural values were averaged before statistical analysis. Texture data were reported as hardness (defined as the maximum shear force, expressed in Newtons), and work (defined as the working shear force, expressed in 10^{-3} J). Texture measures from cheeses collected in the first year (349 samples) were excluded from the database because a different instrument was used.

Color Measurements

Color was measured at 3 different positions on the surface of all cheese samples about 1 h after sample cutting using a portable spectrocolorimeter (CM 508, Minolta Co. Ltd., Osaka, Japan). Before the analysis, the instrument was standardized using a white calibration cap (CM-A70, Minolta Co. Ltd.). Color traits were expressed according to the Commission Internationale de l'Éclairage colorimetric system (CIE L*a*b*), using primary illuminant D65 (standard daylight) with a 10° observer. Briefly, the L* value is the lightness coefficient (ranging from black = 0 to white = 100), the a* value indicates the position on the green (−)-to-red (+) axis, b* the position on the blue (−)-to-yellow (+) axis. The instrument also directly calculates C* (chroma) and h° (hue angle) values: C* is calculated as $(a^{*2} + b^{*2})^{1/2}$, h° is calculated from the arctangent of b*/a*, and together they represent the polar coordinates of the cheese in the color plane in which a* and b* represent the Cartesian (rectangular) coordinates. The 3 acquisitions were averaged before statistical analysis.

Sources of Variation and Repeatability of the Chemical and Physical Traits of Cheese

To assess the relative importance of different sources of variation (cheese category, cheese within category, sample within cheese, year of sampling) and the repeatability of the chemical components of cheese (moisture, ash, protein, lipids, and WSN), 30 cheeses from 15 categories (2 cheeses per category, randomly selected) were sampled twice. Each sample (2 per cheese) was analyzed in duplicate (2 replicates), giving a total of 120 analyses per each chemical component. As specific cheese traits were analyzed in duplicate (texture traits) or triplicate (pH and color traits), the relative importance of each different source of variation was assessed on all cheeses.

These data were analyzed using a mixed model that included the random effects of cheese category, cheese (within cheese category), sample (within cheese, only for chemical components), year of sampling (only for pH and physical traits), and residual (replicate within sample). Repeatability (%) was calculated as the total variance minus the sample and residual variance for chemical composition, and the total variance minus the residual variance for pH, texture, and color traits and expressed as percentage of total variance.

Spectra Collection

Spectra were recorded from all cheeses using 3 different instruments, whose main characteristics are summarized in Table 2.

The portable NIR spectrophotometer working in reflectance mode (**VisNIRS-R**; LabSpec2500, Qualityspec Pro, Analytical Spectral Devices Inc., Boulder, CO) scans the region from 350 (visible area) to 1,830 (NIR) nm, at data-point intervals of 1 nm and is coupled to a high-intensity probe (Analytical Spectral Devices Inc.). Spectra were collected by passing the scanning head of the fiber-optic contact probe (10-mm diameter) over different sites on the freshly cut cheese surface, without any sampling, grinding, or other preparation.

The benchtop NIRS instrument working in reflectance mode (**NIRS-R**; NIRSystem5000, Foss Electric A/S, Hillerød, Denmark) scans the region from 1,100 to 2,498 nm (NIR), at data-point intervals of 2 nm; before spectra collection, cheese samples were ground with a Retsch Grindomix (Retsch GmbH, Haan, Germany) at 4,000 rpm for 10 s, and spectra were collected on about 2 g of grated cheese placed in a 50-mm-diameter ring cup.

The benchtop NIRS instrument working in transmittance mode (**NIRS-T**; FoodScan Dairy Analyzer, Foss Electric A/S), with a moving monochromator, scans the region from 850 to 1,048 nm (NIR), at data-point intervals of 2 nm. In this spectrum interval, the principal types of vibrations are N-H second (∼850 nm) and third (∼1,000 nm) overtones, characteristic of proteins; C-H third overtone (∼900 nm), characteristic of lipids; and O-H second overtone (∼950 nm), typical of water (Osborne, 2006). Spectra were collected from ground cheese (about 150 g) placed in a 90-mm-diameter Petri dish, and milling was as for the NIRS-R instrument.

Analytical Data and Spectra Editing

Values of the chemical and physical traits outside the interval of the mean ± 3 standard deviations (**SD**) were excluded as outliers, and the total number of samples used for calibration procedures (reference data) is reported in Table 1. Outlier spectra of samples with reference data were detected by centering the absorbance

values of each wave in the spectra and standardizing them to a null mean and a unit sample variance within year. Samples exhibiting a large spectral distance (i.e., a global Mahalanobis distance >3) were considered outliers and discarded from the calibration analysis. Mahalanobis distance represents the multidimensional distance, in SD units, of the spectrum of a cheese sample from the centroid of the population of all cheese spectra. It is calculated on the absorbance values of each wavelength, taking into account their covariance matrix. It corresponds to the standard Euclidean distance in the transformed space. The mean absorbance (±SD) recorded for each wavelength on all cheeses with the 3 instruments is depicted in Figure 1, together with the different sections of the spectra used for comparing different instruments on common spectral ranges.

Chemometric Models

A total of 8 calibration models were compared for all the chemical and physical traits. First, the full spectral interval of each instrument was used to perform the calibration equations (3 models). As shown in Figure 1, to compare the instruments with the same spectral intervals and assess the relative importance of different wavelength intervals of the spectrum within instrument, the VisNIRS-R spectrum was divided into 3 portions (3 models) as follows: (1) 350 to 850 nm (mainly visible electromagnetic radiation), (2) 850 to 1,050 nm (mainly third overtone of NIR), and (3) 1,100 to 1,830 nm (mainly second and first overtone of NIR), and the NIRS-R spectrum was divided into 2 portions (2 models) as follows: (1) 1,100 to 1,830 nm (mainly second and first overtone of NIR) and (2) 1,830 to 2,498 nm (mainly combination bands; Table 2). The whole spectrum of NIRS-T corresponds to the second portion of VisNIRS-R (850–1,050 nm, mainly third overtone of NIR).

Chemometric Procedure

Given that the accuracy of the prediction equations is affected by several factors, including the procedure used to develop the calibration equations (Karoui et al., 2010), we decided to use a common calibration procedure to reduce the sources of variation. We therefore extracted the raw spectra from the instrument databases and used a Bayesian approach, which in a previ-

Table 2. Main characteristics of the infrared spectrometers[1] used for predicting chemical composition, pH, texture, and color traits of cheeses

Item	VisNIRS-R	NIRS-R	NIRS-T
Instrument			
Denomination	LabSpec2500	NIRSystem5000	FoodScan
Producer	ASD Inc.	Foss Electric A/S	Foss Electric A/S
City	Boulder, CO	Hillerød	Hillerød
Country	USA	Denmark	Denmark
Characteristics			
Type	Portable	Laboratory benchtop	Laboratory benchtop
Method	Reflectance (R)	Reflectance	Transmittance (T)
Light source	Halogen	Halogen	Halogen
Light detection	External probe	Inner chamber	Inner chamber
Detector type	Diode array (Si, InGaAs)	Monochromator (PbS)	Monochromator (Si)
Spectrum acquisition			
Replicates per spectrum	10	32	16
Spectra collected	3	1	1
Absorbance (A) calculation	$A = \log(1/R)$	$A = \log(1/R)$	$A = \log(1/T)$
Sample preparation	None	Grinding	Grinding
Whole spectrum			
Wavelength range (nm)	350–1,830	1,100–2,498	850–1,048
Data point interval (nm)	1	2	2
Data points per sample	1,481	700	100
Spectrum intervals compared			
Visible, 350–850 nm	Yes	—	—
NIR, 850–1,050 nm	Yes	—	Yes
NIR, 1,100–1,830 nm	Yes	Yes	—
NIR, 1,830–2,498 nm	—	Yes	—
Number of cheeses used			
Chemical composition	193	195	190
pH	955	949	942
Texture traits	579	577	572
Color traits	988	1,021	1,002

[1]VisNIRS-R = portable near-infrared (NIR) spectrophotometer (NIRS) working in reflectance mode; NIRS-R = benchtop NIRS working in reflectance mode; NIRS-T = benchtop NIRS working in transmittance mode.

ous study on milk trait predictions (Ferragina et al., 2015) was found to have greater predictive power than multivariate methods (i.e., PLS and modified PLS). All the chemometric models were implemented in the R environment (R Core Team, 2013). The Bayes B model was fitted to the spectral data, without any information on cheese, producer, category, group, or year.

This model has a mixed prior density assigned to the effects (i.e., point of mass at zero and scaled-t) and is a variable selection method based on BGLR. It is used for calibration models and was developed in the BGLR R package (de los Campos and Pérez-Rodriguez, 2014); the procedure is described in detail in Ferragina et al. (2015). Some detail on the statistical approach and the simplified scripts showing how the predictive equations for the Bayes B model can be implemented in BGLR can be found in the Appendix, whereas the data files relative to the cheese spectra and the reference data are freely available on request from the corresponding author of this paper.

Assessing Predictive Accuracy

A training-testing procedure was used to assess prediction accuracy. The data set was split into a training set (**TRN**) and a testing set (**TST**): 20% of all samples were randomly assigned to the TST as a validation set, and the remaining samples were used as a calibration set in the TRN. The TRN-TST procedure was repeated 10 times, changing the TRN and TST at each replica. We evaluated the prediction performance of the 8 Bayes B models applied to each of the 13 chemical and physical traits of the cheeses on the basis of their determination coefficient of calibration (R^2_{CAL}), and of validation (R^2_{VAL}; where R^2 values between 0.50 and 0.65 indicate low prediction accuracy; values between 0.66 and 0.81 reveal approximate quantitative predictions; values between 0.82 and 0.90 indicate good predictions; values ≥ 0.91 are considered to be excellent; Williams, 2003), and the root mean squared error of validation ($RMSE_{VAL}$); each value was the average of the 10 TRN-TST replicas.

Evaluation of Cheese Categories with Biased Infrared Predictions

To take into account the considerable heterogeneity of the cheese categories, the difference between the predicted (VisNIRS-R, entire spectrum) and measured values of each cheese was calculated for chemical and color traits and then analyzed using the GLM procedure (SAS Institute Inc., Cary, NC), which included only cheese category as a fixed factor. The expected value of the difference between predicted and measured values of each cheese is 0.00. If the least squares means of a cheese category for a predicted trait was significantly different from 0.00 ($P < 0.05$), the category was considered biased for that chemical or physical trait,
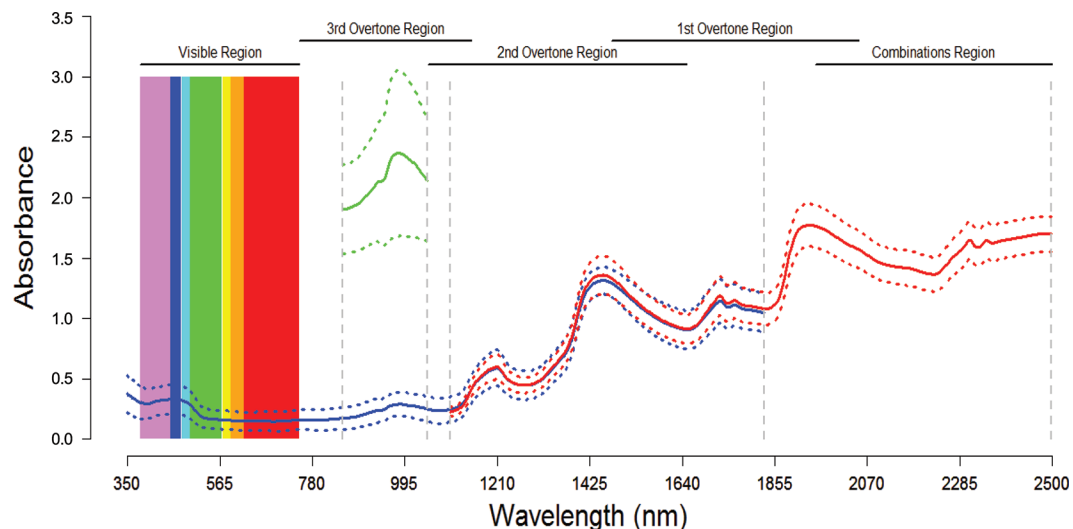

Figure 1. Absorbance spectra (the solid lines represent the average absorbance and the broken lines the mean ± 1 SD) of VisNIRS-R (blue lines, 988 cheeses), NIRS-R (red lines, 1,021 cheeses), and NIRS-T (green lines, 1,002 cheeses). The vertical dotted gray lines separate the spectral intervals considered when comparing the instruments on similar wavelength ranges. VisNIRS-R = portable near-infrared spectrophotometer (NIRS) working in reflectance mode; NIRS-R = benchtop NIRS working in reflectance mode; NIRS-T = benchtop NIRS working in transmittance mode.

Table 3. Source of variation and repeatability of the cheese characteristics determined according to the gold standard methods (chemical composition, pH, texture, and color traits) and used as reference values for infrared predictions

Item	Sum of variances	Source of variation (% of total variance)					RMSE[1]	Repeatability[2] (%)
		Cheese category[3]	Cheese (category)[4]	Sample (cheese)	Year	Residual		
Chemical composition								
DM	212.88	95.0	4.4	0.6	—	0.1	0.32	99.3
Ash	1.96	91.3	6.7	1.7	—	0.3	0.55	98.0
Protein	33.02	94.4	1.8	2.7	—	1.1	1.05	96.2
Lipids	58.86	81.7	16.0	0.6	—	1.6	1.26	97.8
WSN[5]	8.10	82.5	10.4	6.3	—	0.8	0.89	92.9
pH	0.17	42.0	34.1	—	20.8	3.0	1.73	97.0
Texture trait								
Hardness	59.78	43.4	38.4	—	2.8	15.4	3.92	84.6
Work	54.80	28.6	35.2	—	15.6	20.6	4.54	79.4
Color trait[6]								
L*	72.24	53.4	36.0	—	2.0	8.7	2.95	91.3
a*	0.89	13.1	71.3	—	0.9	14.6	3.82	85.4
b*	10.91	27.6	60.1	—	3.8	8.4	2.90	91.6
C*	10.62	27.6	60.0	—	3.7	8.6	2.93	91.4
h°	26.47	19.0	67.4	—	2.7	11.0	3.32	89.0

[1]RMSE = root mean squared error.

[2]The repeatability value corresponds to the maximum R^2 value achievable with calibrations.

[3]Cheese category = type of cheese.

[4]Cheese (category) = cheese sample from different cheese producers within cheese category.

[5]Water-soluble nitrogen.

[6]L* = lightness; a* = bluish-green to red-purple hue component; b* = yellow to blue hue component; C* = chroma; h° = hue angle (0° = red-purple, 90° = yellow, 180° = bluish-green, 270° = blue).

underestimated if the sign of the least squares means was negative, and overestimated if positive.

RESULTS

Sources of Variation and Repeatability of Cheese Characteristics

Results from the ANOVA (mixed model) used for quantifying the sources of variation and repeatability of the cheese characteristics analyzed according to the gold standard methods and used as reference for infrared calibrations are reported in Table 3. Category of cheese was the most important source of variation for all chemical components, accounting for about 82% of total variance for lipids and WSN and 91 to 95% for ash, protein, and DM (Table 3). Much less variability of all traits existed among cheeses of the same category, from 2 to 10% of total variance, with the exception of lipids (16%). Different results were observed for pH and physical traits. Variability among individual cheeses within each category was either almost similar (pH, texture traits, and lightness) or even greater (other color traits) than among cheese categories. Year of sampling represented an appreciable source of variation only for pH, which had a very small total variance, and work, which was much more affected by instrument

setting than the hardness trait. In the case of chemical components, results from 2 samples with 2 replicates of each cheese allowed us to distinguish the proportion of variability due to sampling (differences between samples of the same cheese) and that due to analysis (residual, between replicates of the same sample). Both sources of variation were low (from 0.1 to 2.7% of total variance) for all chemical traits, except for the sampling variability of WSN (6.3%). This means that the overall repeatability, taking into account both sampling and instrumental variability, was greater than 96% for all chemical components, with the exception of WSN (Table 3). The residual variability, combining sampling and instrumental variability, was relatively low for pH (3%) and larger for color (from 8 to 15%) and texture traits (15 and 21%) so that repeatability was in the range of about 80 to 97% (Table 3).

Prediction of the Chemical Composition of Cheese

The results obtained from the calibrations for chemical components using different instruments and spectral ranges are given in Table 4. In general, moisture (R^2_{VAL} = from 0.77 to 0.96), protein (from 0.73 to 0.91), and lipids (from 0.49 to 0.85) were better fitted than ash (from 0.63 to 0.68), WSN (from 0.37 to 0.55), and pH (from 0.12 to 0.24). When results from

different instruments (entire spectrum) were compared, the VisNIRS-R instrument was the most accurate in calibration (R^2_{CAL}) and validation (R^2_{VAL}) procedures, except for WSN. This trait obtained a slightly greater R^2_{CAL} value with NIRS-R, although the R^2_{VAL} and root mean squared error of validation values were slightly better with VisNIRS-R. Predictive performances of partial spectral intervals within the same instrument were, in general, similar and comparable to those using the entire spectrum, excluding few exceptions with VisNIRS-R. In fact, the visible range (350–850 nm) yielded a lower R^2_{VAL} than the other (NIR) intervals and the entire spectrum for all chemical traits, except for ash and WSN (the 850–1,050 nm interval yielded the lowest R^2_{VAL} for the latter; Table 4).

The ranking of instrument performances with the common spectral range intervals (1,100–1,830 nm for comparison of VisNIRS-R and NIRS-R, and 850–1,050 nm for comparison of VisNIRS-R and NIRS-T) was similar to that obtained using the full spectrum (VisNIRS-R > NIRS-R > NIRS-T).

Prediction of the Physical Traits of Cheeses

In Table 5 are reported the results obtained from calibrations for cheese texture and color traits. Hardness had a low fit (R^2_{VAL} = from 0.37 to 0.46), although it was better than work (R^2_{VAL} = from 0.13 to 0.18). The differences among and within instruments at different spectral intervals were low for these 2 texture traits.

The lightness color trait was predicted with moderate accuracy (R^2_{VAL} = from 0.57 to 0.75) by all the instruments and all the spectral ranges, with the best result from VisNIRS-R, followed by NIRS-R and NIRS-T. Accuracy of prediction of all the other cheese color traits by all the instruments and spectral intervals was low (R^2_{VAL} = from 0.01 to 0.20), with the (expected) exception of the VisNIRS-R predictions and, particularly, its visible interval (350 to 850 nm) in the spectrum (R^2_{VAL} = from 0.47 to 0.61).

Estimation of Prediction Biases for Some Cheese Categories

Descriptive statistics and least squares means of the biases (where the average of the residuals of the cheeses of a given category differed significantly from the expected 0.00 value) in the chemical and color traits of under- and overestimated cheese categories are summarized in Table 6. Cheese pH and texture traits have not been reported here because of their low prediction accuracy. We found a moderate number of biased categories for chemical and color traits (13 and 9%,

respectively). In particular, the caprine cheese types showed some biases in prediction of chemical composition (goat-sour cheeses) and in color traits (goat-rennet cheeses). Mozzarella cheese also showed some biases for protein (underestimated), lipids, and some color traits (overestimated). Also other fresh cheeses (including Casatella Trevigiana PDO) showed some bias and, among traditional cheeses, the Morlacco del Grappa category. Other mountain PDO cheeses were characterized by biased chemical components estimation, and moldy rind and blue cheeses showed an underestimation of WSN. Last, biases (WSN and color traits) were observed for some cheese categories with inclusion of spices (pepper, chili cheeses, Table 6).

DISCUSSION

Prediction of Cheese Quality Traits

To our knowledge, no previous studies have investigated the use of different instruments and spectral intervals in the visible and NIR ranges to predict the chemical and physical traits of different cheeses of various categories (Table 1). Discussion will, therefore, focus mostly on studies on specific cheese types and techniques.

The only large study was carried out by Lucas et al. (2008) on 445 samples of 4 types of cheese (Abondance, Tomme de Savoie, Cantal cows' milk cheeses, and Rocamadour goat cheese) to predict (using modified PLS regression) several cheese characteristics using a laboratory benchtop VisNIR spectrometer (400–2,500 nm spectrum range). They obtained R^2_{VAL} values very close to ours for predicting moisture (0.97 vs. 0.96) and fat (0.90 vs. 0.85), but their root mean squared error (**RMSE**) values were lower than ours, potentially due to lower variability in their sample set. Unlike our study, Lucas et al. (2008) did not quantify possible specific biases for different cheese types. In a study on 91 Emmental cheese samples, Karoui et al. (2006) obtained a greater cross-validated R^2 (0.94) than we found for lipids (0.85 with NIR, 1,000–2,500 nm spectrum range) and for WSN predictions (0.88 vs. 0.53) but a similar value to ours for protein prediction (0.86 vs. 0.88) and a lower value for salt prediction (0.50) than ours for total ash prediction (0.75). The greater values of cross-validated R^2 and the lower RMSE values found by Karoui et al. (2006) could be due to the fact that the samples analyzed were highly homogeneous (one category, Emmental). Anyway, the main disadvantage of NIR spectroscopy is its weak sensitivity to minor constituents such as salt and WSN (Berzaghi and Riovanto, 2009), so the low performances for WSN were expected. Actu-

Table 4. Results of calibration and validation for chemical composition on the spectra obtained with the 3 instruments[1] and different spectral intervals[2]

Chemical composition	Spectral interval	Range (nm)	R^2_{CAL}[3]	R^2_{VAL}[4]	$RMSE_{VAL}$[5]
Moisture					
VisNIRS-R	Whole spectrum	350–1,830	0.97	0.96	2.10
VisNIRS-R	Visible light	350–850	0.88	0.85	4.15
VisNIRS-R	NIR, third overtone (OT)	850–1,050	0.96	0.96	2.30
VisNIRS-R	NIR, second and first OT	1,100–1,830	0.97	**0.96**	**2.00**
NIRS-R	Whole spectrum	1,100–2,498	0.85	0.83	4.51
NIRS-R	NIR, second and first OT	1,100–1,830	0.83	0.84	4.39
NIRS-R	NIR, combination bands	1,830–2,498	0.83	0.81	4.79
NIRS-T	Whole spectrum (third OT)	850–1,050	0.79	0.77	5.26
Protein					
VisNIRS-R	Whole spectrum	350–1,830	0.91	0.88	1.77
VisNIRS-R	Visible light	350–850	0.87	0.84	2.03
VisNIRS-R	NIR, third OT	850–1,050	0.92	**0.91**	**1.57**
VisNIRS-R	NIR, second and first OT	1,100–1,830	0.93	0.91	1.59
NIRS-R	Whole spectrum	1,100–2,498	0.82	0.81	2.25
NIRS-R	NIR, second and first OT	1,100–1,830	0.79	0.76	2.51
NIRS-R	NIR, combination bands	1,830–2,498	0.80	0.79	2.38
NIRS-T	Whole spectrum (third OT)	850–1,050	0.75	0.73	2.62
Lipids					
VisNIRS-R	Whole spectrum	350–1,830	0.87	**0.85**	**2.03**
VisNIRS-R	Visible light	350–850	0.58	0.49	3.73
VisNIRS-R	NIR, third OT	850–1,050	0.85	**0.85**	**2.03**
VisNIRS-R	NIR, second and first OT	1,100–1,830	0.86	**0.85**	**2.03**
NIRS-R	Whole spectrum	1,100–2,498	0.70	0.67	3.20
NIRS-R	NIR, second and first OT	1,100–1,830	0.69	0.69	3.11
NIRS-R	NIR, combination bands	1,830–2,498	0.68	0.63	3.42
NIRS-T	Whole spectrum (third OT)	850–1,050	0.66	0.64	3.28
Water-soluble nitrogen					
VisNIRS-R	Whole spectrum	350–1,830	0.53	0.53	**1.94**
VisNIRS-R	Visible light	350–850	0.49	0.53	**1.94**
VisNIRS-R	NIR, third OT	850–1,050	0.32	0.37	2.26
VisNIRS-R	NIR, second and first OT	1,100–1,830	0.49	0.51	1.98
NIRS-R	Whole spectrum	1,100–2,498	0.57	0.52	2.12
NIRS-R	NIR, second and first OT	1,100–1,830	0.46	0.44	2.36
NIRS-R	NIR, combination bands	1,830–2,498	0.54	**0.55**	2.04
NIRS-T	Whole spectrum (third OT)	850–1,050	0.42	0.47	2.05
Ash					
VisNIRS-R	Whole spectrum	350–1,830	0.75	0.67	0.63
VisNIRS-R	Visible light	350–850	0.71	0.67	0.63
VisNIRS-R	NIR, third OT	850–1,050	0.73	0.67	0.63
VisNIRS-R	NIR, second and first OT	1,100–1,830	0.75	**0.68**	**0.62**
NIRS-R	Whole spectrum	1,100–2,498	0.65	0.68	0.65
NIRS-R	NIR, second and first OT	1,100–1,830	0.58	0.63	0.70
NIRS-R	NIR, combination bands	1,830–2,498	0.64	0.68	0.65
NIRS-T	Whole spectrum (third OT)	850–1,050	0.66	0.65	0.66
pH					
VisNIRS-R	Whole spectrum	350–1,830	0.21	0.24	0.26
VisNIRS-R	Visible light	350–850	0.16	0.19	0.27
VisNIRS-R	NIR, third OT	850–1,050	0.17	0.19	0.27
VisNIRS-R	NIR, second and first OT	1,100–1,830	0.20	0.24	0.26
NIRS-R	Whole spectrum	1,100–2,498	0.21	0.18	0.27
NIRS-R	NIR, second and first OT	1,100–1,830	0.18	0.17	0.28
NIRS-R	NIR, combination bands	1,830–2,498	0.19	0.17	0.28
NIRS-T	Whole spectrum (third OT)	850–1,050	0.11	0.12	0.29

[1]VisNIRS-R = portable near-infrared (NIR) spectrophotometer (NIRS) working in reflectance mode; NIRS-R = benchtop NIRS working in reflectance mode; NIRS-T = benchtop NIRS working in transmittance mode.

[2]The greatest R^2_{VAL} and lowest $RMSE_{VAL}$ for each predicted trait are in bold only if the R^2_{VAL} is >0.50.

[3]R^2_{CAL} = average from 10 replications of the R^2 of calibration.

[4]R^2_{VAL} = average from 10 replications of the R^2 of validation.

[5]$RMSE_{VAL}$ = average from 10 replications of the root mean squared error of validation.

Table 5. Results of calibration and validation for texture and color traits of cheeses on the spectra obtained with the 3 instruments[1] and different spectral intervals[2]

Item	Spectral interval	Range (nm)	R^2_{CAL}[3]	R^2_{VAL}[4]	$RMSE_{VAL}$[5]
Texture trait					
Hardness					
VisNIRS-R	Whole spectrum	350–1,830	0.47	0.46	5.03
VisNIRS-R	Visible light	350–850	0.44	0.44	5.13
VisNIRS-R	NIR, third overtone (OT)	850–1,050	0.45	0.45	5.09
VisNIRS-R	NIR, second and first OT	1,100–1,830	0.46	0.46	5.08
NIRS-R	Whole spectrum	1,100–2,498	0.45	0.43	5.02
NIRS-R	NIR, second and first OT	1,100–1,830	0.43	0.43	5.03
NIRS-R	NIR, combination bands	1,830–2,498	0.44	0.42	5.09
NIRS-T	Whole spectrum (third OT)	850–1,050	0.37	0.37	5.41
Work					
VisNIRS-R	Whole spectrum	350–1,830	0.17	0.17	6.77
VisNIRS-R	Visible light	350–850	0.15	0.17	6.79
VisNIRS-R	NIR, third OT	850–1,050	0.17	0.18	6.75
VisNIRS-R	NIR, second and first OT	1,100–1,830	0.18	0.19	6.73
NIRS-R	Whole spectrum	1,100–2,498	0.16	0.18	6.56
NIRS-R	NIR, second and first OT	1,100–1,830	0.14	0.17	6.61
NIRS-R	NIR, combination bands	1,830–2,498	0.15	0.18	6.57
NIRS-T	Whole spectrum (third OT)	850–1,050	0.15	0.13	6.97
Color trait					
Lightness (L*)					
VisNIRS-R	Whole spectrum	350–1,830	0.72	**0.75**	**3.93**
VisNIRS-R	Visible light	350–850	0.71	0.74	4.02
VisNIRS-R	NIR, third OT	850–1,050	0.70	0.73	4.10
VisNIRS-R	NIR, second and first OT	1,100–1,830	0.71	0.74	4.02
NIRS-R	Whole spectrum	1,100–2,498	0.70	0.66	4.51
NIRS-R	NIR, second and first OT	1,100–1,830	0.70	0.67	4.49
NIRS-R	NIR, combination bands	1,830–2,498	0.67	0.64	4.62
NIRS-T	Whole spectrum (third OT)	850–1,050	0.62	0.57	5.05
Green–red (a*)					
VisNIRS-R	Whole spectrum	350–1,830	0.52	**0.50**	**0.61**
VisNIRS-R	Visible light	350–850	0.51	**0.50**	**0.61**
VisNIRS-R	NIR, third OT	850–1,050	0.02	0.01	0.86
VisNIRS-R	NIR, second and first OT	1,100–1,830	0.03	0.02	0.86
NIRS-R	Whole spectrum	1,100–2,498	0.03	0.01	0.83
NIRS-R	NIR, second and first OT	1,100–1,830	0.04	0.02	0.83
NIRS-R	NIR, combination bands	1,830–2,498	0.03	0.02	0.83
NIRS-T	Whole spectrum (third OT)	850–1,050	0.01	0.01	0.82
Yellow–blue (b*)					
VisNIRS-R	Whole spectrum	350–1,830	0.65	**0.61**	**1.82**
VisNIRS-R	Visible light	350–850	0.63	0.60	1.85
VisNIRS-R	NIR, third OT	850–1,050	0.18	0.16	2.67
VisNIRS-R	NIR, second and first OT	1,100–1,830	0.24	0.22	2.58
NIRS-R	Whole spectrum	1,100–2,498	0.20	0.15	2.78
NIRS-R	NIR, second and first OT	1,100–1,830	0.18	0.14	2.79
NIRS-R	NIR, combination bands	1,830–2,498	0.21	0.17	2.73
NIRS-T	Whole spectrum (third OT)	850–1,050	0.08	0.06	2.85
Chrome (C*)					
VisNIRS-R	Whole spectrum	350–1,830	0.65	**0.61**	**1.79**
VisNIRS-R	Visible light	350–850	0.63	0.60	1.82
VisNIRS-R	NIR, third OT	850–1,050	0.18	0.16	2.63
VisNIRS-R	NIR, second and first OT	1,100–1,830	0.24	0.22	2.53
NIRS-R	Whole spectrum	1,100–2,498	0.20	0.15	2.72
NIRS-R	NIR, second and first OT	1,100–1,830	0.18	0.14	2.75
NIRS-R	NIR, combination bands	1,830–2,498	0.21	0.17	2.69
NIRS-T	Whole spectrum (third OT)	850–1,050	0.08	0.03	4.71
Hue angle (h°)					
VisNIRS-R	Whole spectrum	350–1,830	0.51	0.47	3.50
VisNIRS-R	Visible light	350–850	0.50	0.47	3.52
VisNIRS-R	NIR, third OT	850–1,050	0.05	0.03	4.76
VisNIRS-R	NIR, second and first OT	1,100–1,830	0.08	0.06	4.66
NIRS-R	Whole spectrum	1,100–2,498	0.09	0.04	4.76
NIRS-R	NIR, second and first OT	1,100–1,830	0.08	0.05	4.76
NIRS-R	NIR, combination bands	1,830–2,498	0.08	0.05	4.75
NIRS-T	Whole spectrum (third OT)	850–1,050	0.04	0.03	4.71

[1]VisNIRS-R = portable near-infrared (NIR) spectrophotometer (NIRS) working in reflectance mode; NIRS-R = benchtop NIRS working in reflectance mode; NIRS-T = benchtop NIRS working in transmittance mode.

[2]The greatest R^2_{VAL} for each predicted trait is in bold only if it is >0.50.

[3]R^2_{CAL} = average from 10 replications of the R^2 of calibration.

[4]R^2_{VAL} = average from 10 replications of the R^2 of validation.

[5]$RMSE_{VAL}$ = average from 10 replications of the root mean squared error of validation.

ally, cheese content of WSN was predicted using NIRS (400–1,100 nm spectrum range) also by Kraggerud et al. (2014) with modest results (R^2_{CV} = 0.42; RMSE of cross validation = 0.12%). Botelho et al. (2013), analyzing 123 Mozzarella samples with a Fourier transform infrared spectrophotometer (1,000–2,500 nm spectrum range), obtained RMSE of prediction very similar to ours (2.2% for moisture and 3.2% for fat content) by using PLS regression. Considering another dairy product, Madalozzo et al. (2015) tested NIRS (1,100–2,500 nm spectrum range) with PLS regression on 19 ricotta samples in duplicate and obtained results for fat and protein predictions very similar to those of Karoui et al. (2006) in terms of cross-validated R^2 and to ours in terms of RMSE. Their prediction of cheese moisture (%) had a lower R^2 (0.76 vs. 0.96) and a greater RMSE (3.3 vs. 2.1) than our prediction. Recently, Margolies and Barbano (2018), determining the composition of 70 Cheddar cheese samples by using MIR transmittance spectroscopy, obtained prediction of moisture and fat content characterized by very similar R^2_{VAL} (0.93 and 0.70, respectively), but lower RMSE of validation, likely because of the much lower variability characterizing a single cheese-type data set.

Our prediction accuracies for the pH (R^2_{VAL} = from 0.12 to 0.24) and texture (R^2_{VAL} = from 0.13 to 0.46) traits were weak (Tables 4 and 5); they were much lower than the maximum achievable R^2 corresponding to the repeatability of reference traits (Table 3). This could be explained by the fact that these measurements are not linked to specific chemical compounds so that no specific absorbance bands are recognizable in the visible and NIR regions, and their predictions can only be indirectly based on correlations with some other chemical traits. Similar results for pH were obtained also by Karoui et al. (2006; R^2 = 0.40) and Kraggerud et al. (2014; R^2 = 0.12) using NIRS, whereas Lucas et al. (2008) obtained slightly better results using VisNIRS. Downey et al. (2005) obtained weak results for cheese firmness with NIRS (400–2,500 nm spectrum range) using PLS regression (R^2 = from 0.25 to 0.41), as in our study.

The adaptability of the model for color traits was not satisfactory in our study (Table 5), with the exception

Table 6. Descriptive statistics (mean ± SD) and least squares means of differences between predicted and measured values (bias) for chemical composition and color traits for the cheese categories significantly deviating from zero (on a total of 37 cheese categories tested for each trait)

		Underestimated		Overestimated	
Trait	Mean ± SD	Bias	Category	Bias	Category
Chemical composition (%)					
Moisture	40.3 ± 10.8	−2.35**	Casatella Trevigiana PDO[1]	+1.28*	Morlacco del Grappa
		−2.18*	Provolone Valpadana PDO piccante		
		−3.03***	Goat-sour cheeses		
Protein	24.1 ± 5.2	−1.22*	Montasio PDO 2–5 mo	+1.65**	Casatella Trevigiana PDO
		−1.43*	Monte Veronese PDO fresh	+2.17***	Fresh cheeses
		−2.33***	Mozzarella cheese	+1.14*	Hard cheeses
Lipids	28.5 ± 5.3	−3.49***	Fresh cheeses	+1.59*	Asiago PDO 4–6 mo
		−2.06*	Stretched-soft cheeses	+1.63*	Mozzarella cheese
				+2.95***	Goat-sour cheeses
Water-soluble nitrogen	6.1 ± 2.9	−4.56***	Moldy rind cheeses	+1.25*	Fresh cheeses
		−3.84***	Blue cheeses	+2.10**	Pepper, chili cheeses
				+1.32*	Goat-sour cheeses
Ash	3.7 ± 1.1	−0.63*	Monte Veronese PDO >12 mo	+0.52*	Casatella Trevigiana PDO
		−0.76***	Morlacco del Grappa	+0.50*	Goat-sour cheeses
Color trait[2]					
L*	76.8 ± 7.8	−1.22*	Fresh cheeses	+1.18*	Asiago PDO pressato
		−4.05*	Beer, vinegar refined cheeses		
a*	−1.0 ± 0.9	−0.25*	Morlacco del Grappa	+0.26*	Goat-rennet cheeses
		−0.17*	Soft cheeses with rind <1 kg		
		−0.36***	Pepper, chili cheeses		
b*	12.4 ± 3.1	−0.72*	Morlacco del Grappa	+1.24**	Mozzarella cheese
		−0.61*	Malga fresco 2–4 mo	+0.70*	Goat-rennet cheeses
C*	12.5 ± 3.1	−0.61*	Malga fresco 2–4 mo	+1.29**	Mozzarella cheese
h°	95.6 ± 4.9	−2.71***	Goat-rennet cheeses	+1.99**	Morlacco del Grappa
		−1.66*	Montasio PDO >10 mo	+1.26*	Pepper, chili cheeses

[1]PDO = protected designation of origin.

[2]L* = lightness; a* = bluish-green to red-purple hue component; b* = yellow–blue hue component; C* = chroma; h° = hue angle (0° = red-purple, 90° = yellow, 180° = bluish-green, 270° = blue).

***$P < 0.001$; **$P < 0.01$; *$P < 0.05$.

of the VisNIRS-R instrument, which yielded moderate results for all color traits (R^2_{VAL} = from 0.47 to 0.75). Lucas et al. (2008) obtained better prediction performances than we did for cheese color traits on Emmental cheese with a VisNIR instrument ($R^2_{CV} \geq 0.93$).

As a general comment, we need to bear in mind that the calibration accuracy obtainable with a secondary technique, such as infrared spectroscopy, cannot be greater than the repeatability of the reference measures, including both sampling and instrumental variability, and also the year of sampling. This means that, looking at our results, the maximum R^2_{VAL} that can be expected for these traits is not 100% but ranges from about 90% for lightness to 64% for work (Table 2). The moderate accuracy of the NIRS predictions also depends heavily on the lower repeatability of the physical properties of cheese.

Comparison of Spectroscopic Instruments

Very few studies can be found in the scientific literature that compare the predictions of given properties of cheese made by different instruments, operating in different ranges of the electromagnetic spectrum (Karoui et al., 2006; Andueza et al., 2013). In these cases, the effects of instrument technology and spectrum range are confounded. On the other hand, no studies have compared different NIR instruments using the same wavelength interval, or different wavelength intervals within the same instrument, as we have attempted in this study.

In accordance with the International Organization for Standardization (ISO, 2015a,b; ISO 11151-1:2015 and 11151-2:2015), the entire electromagnetic spectral area analyzed in the present study with the 3 instruments (wavelengths 350 to 2,498 nm) covers the visible region (Vis, 350 to 750, corresponding to a wavenumber range between about 28,000 and 13,000 $\times$ cm^{-1}) and the near-infrared region (NIR, 750 to 2,500 nm, corresponding to a wavenumber range between 13,000 and 4,000 $\times$ cm^{-1}).

Some studies have compared NIRS with MIRS applied to cheese. Analyzing 90 Camembert-type experimental cheeses by MIRS, Martín del Campo et al. (2007) obtained intermediate results with respect to ours with the VisNIRS and with the 2 NIRS instruments. Mid-infrared spectroscopy seems to yield better results when used for calibrating fat (Margolies and Barbano, 2018), WSN (Kraggerud et al., 2014), pH (Karoui et al., 2006; Martín del Campo et al., 2007), and cheese hardness (Fagan et al., 2007a) but not cheese moisture (Margolies and Barbano, 2018). Again, when results are obtained from a single homogeneous category of

cheese, they may refer mainly to chemical composition and be specific to that cheese type.

Comparing the 3 instruments with their full spectrum range, we found that the predictive performance of the VisNIRS-R was always better than the other 2 benchtop NIR instruments and that, of the latter, the NIR-R performed better than the NIR-T, with a few exceptions.

The breadth of the spectral range analyzed and the number of waves tested are often considered important sources of variation (Liu et al., 2014). In our study, the ranking of the 3 instruments was in line with the ranking of the number of individual waves analyzed (1,481 for VisNIRS-R, 700 for NIR-R, and 100 for NIR-T, respectively). However, we also compared the common intervals of the cheese spectrum of the 3 different instruments. In particular, we were able to compare VisNIRS-R and NIR-R using, in both cases, the central portion of the NIR (wavelength 1,100 to 1,830 nm), although the differences in predictive accuracy remained largely unchanged among traits (Tables 4 and 5). This means that the differences are probably due to the different instruments adopted than to the spectral interval. The light detection system, the detectors, and the sample preparation, in particular, may have contributed to the dissimilarities observed (Table 2).

We were able to directly compare also the VisNIRS-R and NIRS-T instruments by developing calibrations on the common interval represented by the lower part of the NIR spectrum (wavelength 850 to 1,050 nm), and even in this case, the differences remained among traits in terms of predictive accuracy (Tables 4 and 5). Furthermore, NIRS-T is equipped with artificial neural network calibrations (chemical components) for various categories of food, cheese included. However, the predictions provided directly by the instrument software (data not shown) were, in general, less accurate than those obtained in this study, which may be due to the wide variability in our cheese categories. In some cases, these predictions assumed values greater than 100% or were negative. For example, when a goat-sour cheese or a blue cheese was analyzed by NIRS-T, the precalibrated prediction gave values greater than 100% for moisture content, and negative values for protein and lipids components (data not shown).

Comparison of the Spectrum Intervals in the Range of Visible and NIR Electromagnetic Radiation

As far as we are aware, the scientific literature contains no comparisons of the prediction abilities of different portions of the visible and NIR spectrum of cheese. The VisNIRS-R instrument allowed us to draw

inferences regarding the capability of the visible portion of the cheese spectra to predict cheese properties with respect to 2 of the 3 portions of the NIR considered in this study (Figure 1). Referring to chemical composition of cheeses, in the case of moisture and protein, the predictions based on this portion of the spectrum were slightly less effective than those obtained from the other NIRS portions of the same instrument. However, they were similar to or even better than those obtained from the NIRS portions of the other 2 instruments (Table 4). In the case of lipids, the visible portion yielded predictions less accurate than all the NIRS portions of the same and other instruments, whereas they were comparable for WSN, ash, and pH (Table 4).

This portion of the spectrum includes a small part of the UV range (350–400 nm), all the visible ranges (400–750 nm), and also a small part of the near-infrared spectrum (750–850 nm). This interval is characterized by mostly electronic excitations or transitions (i.e., oxidation states). In this part, functional groups with high electron density, such as the carbonyl and nitro groups with double, triple, or conjugated double bonds, are strong absorbers of light (chromophores). Macro components, such as proteins and lipids, usually contain chromophores, whereas inorganic compounds (i.e., ash) are subject to transitions states (Hlavác, 2013). These are probably the reasons why the visible spectrum interval was able to predict the chemical composition of cheeses with sufficient accuracy.

In the case of physical traits, the visible portion of the cheese spectra is at least as effective as the NIRS portions in the case of hardness, work, and lightness and much more effective for the other color traits (a*, b*, C*, and H*, see Table 5). This was expected because this portion includes the 700 to 635 nm range corresponding to red, 635 to 590 nm to orange, 590 to 560 nm to yellow, 560 to 490 nm to green, and 490 to 450 nm to blue (Jha, 2010).

The correlation between absorbance of some specific wavelength of the spectrum and the chemical bonds and functional groups of milk and dairy products is well known for the MIR electromagnetic radiation, especially when the Fourier transform is used (Bittante and Cecchinato, 2013). In the case of the NIR, the correlation is broader. The first portion of the NIRS (850 to 1,050 nm) is the only one analyzed by the NIRS-T, which is specialized in the prediction of the chemical composition of food, and by VisNIRS-R instruments. In fact, this spectral range includes mainly bands assigned to the third overtone and some signals of great interest for food composition (C-H, O-H, and N-H; Osborne, 2006; Lucas et al., 2008). The second portion of the NIRS (1,050 to 1,830 nm), analyzed by the VisNIRS-R and NIRS-R instruments, includes bands assigned

mainly to the second overtone (C-H and O-H, till 1,400 nm) and first overtone (N-H, C-H, and C = O), according to Osborne (2006), Pi et al. (2009), and Lenart et al. (2012), respectively. The third portion of NIRS (1,830 to 2,498 nm) is recognized to be a fingerprint area because of the presence of different bands not clearly related to a specific bond and combination bands. Also in this case, N-H, C-H and O-H bonds are identifiable (Karoui et al., 2006; Madalozzo et al., 2015). The presence of signals related to moisture, fat, and protein on all the 3 portions of the NIRS considered explains why, within instrument, the different portions of NIRS considered (first and second in the case of VisNIRS-R and second and third in the case of NIRS-R) yielded predictions of the chemical composition of cheese similar to each other and also to the results obtained with the entire spectrum. Moreover, it confirms that the differences observed among instruments are more related to technical feature than to spectral intervals used.

Possible Bias in the Predictions of Some Cheese Categories

In this study, we did not intend to compare the different categories of cheese in relation to the predictability of their physical and chemical characteristics by a general cheese calibration, nor to develop cheese category-specific calibrations, because, of course, this would require a much greater number of cheeses to be sampled and analyzed. Nevertheless, we looked at the effect of category (Table 6) on the differences between predicted and measured values (errors of prediction) in the best performing model (VisNIRS-R, entire spectrum) to see whether any biases (under- or overpredictions) affected any cheese categories. It is indeed recognized that cheese components depend on the composition of the original milk and on different conditions of milk production and processing (Lucas et al., 2005), which actually determine cheese category.

As evidenced in Table 6, specific categories (i.e., moldy rind cheeses, blue cheeses) would require particular consideration, because their bacteria modified the composition, especially the amount of WSN. These categories, together with goat-sour cheeses, had very different results from the other common cheeses, also in terms of pH (data not shown). We can suppose that, in the case of chemical composition, the high number of biased predictions of goat-sour cheeses is ascribable to the type of coagulation and presence of bacteria and not to species. This hypothesis is confirmed if we compare goat-sour cheeses with goat-rennet cheese; the latter did not exhibit any bias for chemical components but only in some color traits. The other 6 categories of cheese exhibiting more than one biased trait (5 traits

for Morlacco cheese; 4 traits for Mozzarella cheese and for fresh cheeses; 3 traits for Casatella Trevigiana PDO and for pepper, chili cheeses; and 2 traits for Fresh Malga cheese) were all referring to fresh cheese types. This means that specific calibration for these types of cheeses would be necessary for an accurate prediction of their composition.

CONCLUSIONS

This work provides confirmation that NIR spectroscopy can be used to predict the chemical composition of a large number of different varieties of cheeses but does not seem to be useful for pH and texture characteristics. Prediction accuracy for color traits was insufficient for the 2 laboratory NIRS instruments, whereas it varied according to the different intervals of the visible NIR spectrum. The use of portable instruments in field conditions on the freshly cut cheese surface, without the need to collect, store, transport, and process cheese samples for laboratory analyses, seems to be particularly attractive. Our results also indicate that instrument technology is much more important than the NIR spectral range in terms of accuracy of prediction equations and that the visible range is useful when predicting color traits, other than lightness. However, ad hoc cheese category calibrations seem to be needed to obtain unbiased and more accurate results, especially for goat, moldy, and fresh cheese categories.

REFERENCES

Andueza, D., C. Agabriel, I. Constant, A. Lucas, and B. Martin. 2013. Using visible or near infrared spectroscopy (NIRS) on cheese to authenticate cow feeding regimes. Food Chem. 141:209–214.

AOAC. 1990. Official Methods of Analysis of the Association of Official Analytical Chemists. 15th ed. Assoc. Off. Anal. Chem. Inc., Arlington, VA.

Berzaghi, P., and R. Riovanto. 2009. Near infrared spectroscopy in animal science production: Principles and applications. Ital. J. Anim. Sci. 8:39–62.

Bittante, G., and A. Cecchinato. 2013. Genetic analysis of the Fourier-transform infrared spectra of bovine milk with emphasis on individual wavelengths related to specific chemical bonds. J. Dairy Sci. 96:5991–6006.

Botelho, B. G., B. A. P. Mendes, and M. M. Sena. 2013. Development and analytical validation of robust near-infrared multivariate calibration models for the quality inspection control of Mozzarella cheese. Food Anal. Methods 6:881–891.

Cattaneo, T. M., and S. E. Holroyd. 2013. New applications of near infrared spectroscopy on dairy products. J. Near Infrared Spectrosc. 21:307–310.

de los Campos, G., J. M. Hickey, R. Pong-Wong, H. D. Daetwyler, and M. P. L. Calus. 2013. Whole genome regression and prediction methods applied to plant and animal breeding. Genetics 193:327–345.

de los Campos, G., and P. Pérez-Rodriguez. 2014. BGLR: Bayesian Generalized Linear Regression. Version 1.0.3. Accessed Jan. 5, 2018. http://cran.r-project.org/web/packages/BGLR/index.html.

Downey, G., E. Sheehan, C. Delahunty, D. O'Callaghan, T. Guinee, and V. Howard. 2005. Prediction of maturity and sensory attri-

butes of Cheddar cheese using near-infrared spectroscopy. Int. Dairy J. 15:701–709.

Fagan, C. C., C. Everard, C. P. O'Donnell, G. Downey, E. M. Sheehan, C. M. Delahunty, and D. J. O'Callaghan. 2007a. Evaluating mid-infrared spectroscopy as a new technique for predicting sensory texture attributes of processed cheese. J. Dairy Sci. 90:1122–1132.

Fagan, C. C., C. P. O' Donnel, D. J. O'Callaghan, G. Downey, E. M. Sheehan, C. M. Delahunty, C. Everard, T. P. Guinee, and V. Howard. 2007b. Application of mid-infrared spectroscopy to the prediction of maturity and sensory texture attributes of Cheddar cheese. J. Food Sci. 72:E130–137.

Ferragina, A., G. de los Campos, A. I. Vazquez, A. Cecchinato, and G. Bittante. 2015. Bayesian regression models outperform partial least squares methods for predicting milk components and technological properties using infrared spectral data. J. Dairy Sci. 98:8133–8151.

González-Martín, M. I., P. Severiano-Pérez, I. Revilla, A. M. Vivar-Quintana, J. M. Hernández-Hierro, C. González-Pérez, and I. A. Lobos-Ortega. 2011. Prediction of sensory attributes of cheese by near-infrared spectroscopy. Food Chem. 127:256–263.

Hlavác, M. 2013. Measurement of tissue optical properties. Master's Thesis. Faculty Electr. Eng, Czech Tech. Univ. Prague, Prague, Czech Republic.

ISO. 2001. Milk and Milk Products—Extraction Methods for Lipids and Liposoluble Compounds. 1st ed. Int. Org. Stand., Geneva, Switzerland.

ISO. 2015a. ISO 11151-1:2015 (Lasers and Laser-Related Equipment—Standard Optical Components—Part 1: Components for the UV, Visible and Near-Infrared Spectral Ranges). Int. Org. Stand., Geneva, Switzerland.

ISO. 2015b. ISO 11151-2:2015 (Lasers and Laser-Related Equipment—Standard Optical Components—Part 2: Components for the Infrared Spectral Range). Int. Org. Stand., Geneva, Switzerland.

ISO-IDF. 2006. ISO 21543:2006. IDF 201:2006 (Milk Products—Guidelines for the Application of Near Infrared Spectrometry). Int. Org. Stand., Geneva, Switzerland; Int. Dairy Fed., Brussels, Belgium.

Jha, S. N. 2010. Nondestructive Evaluation of Food Quality. Vol. 375. Springer, Heidelberg, Germany.

Karoui, R., G. Downey, and C. Blecker. 2010. Mid-infrared spectroscopy coupled with chemometrics: A tool for the analysis of intact food systems and the exploration of their molecular structure–quality relationships—A review. Chem. Rev. 110:6144–6168.

Karoui, R., A. M. Mouazen, É. Dufour, L. Pillonel, E. Schaller, J. De Baerdemaeker, and J. Bosset. 2006. Chemical characterisation of European Emmental cheeses by near infrared spectroscopy using chemometric tools. Int. Dairy J. 16:1211–1217.

Kraggerud, H., T. Næs, and R. K. Abrahamsen. 2014. Prediction of sensory quality of cheese during ripening from chemical and spectroscopy measurements. Int. Dairy J. 34:6–18.

Lee, S., I. Jeon, and L. Harbers. 1997. Near infrared reflectance spectroscopy for rapid analysis of curds during Cheddar cheese making. J. Food Sci. 62:53–56.

Lénart, J., T. Szigedi, M. Dernovics, and M. Fodor. 2012. Application of FT-NIR spectroscopy on the determination of the fat and protein contents of lyophilized cheeses. Acta Aliment. 41:351–362.

Liu, D., D. W. Sun, and X. A. Zeng. 2014. Recent advances in wavelength selection techniques for hyperspectral image processing in the food industry. Food Bioprocess Technol. 7:307–323.

Lucas, A., D. Andueza, E. Rock, and B. Martin. 2008. Prediction of dry matter, fat, pH, vitamins, minerals, carotenoids, total antioxidant capacity, and color in fresh and freeze-dried cheeses by visible-near-infrared reflectance spectroscopy. J. Agric. Food Chem. 56:6801–6808.

Lucas, A., E. Rock, J.-F. Chamba, I. Verdier-Metz, P. Brachet, and J.-B. Coulon. 2005. Respective effects of milk composition and the cheese-making process on cheese compositional variability in components of nutritional interest. Lait 86:21–41.

Madalozzo, E. S., E. Sauer, and N. Nagata. 2015. Determination of fat, protein and moisture in ricotta cheese by near infrared

spectroscopy and multivariate calibration. J. Food Sci. Technol. 52:1649–1655.

Margolies, B. J., and D. M. Barbano. 2018. Determination of fat, protein, moisture, and salt content of Cheddar cheese using midinfrared transmittance spectroscopy. J. Dairy Sci. 101:924–933.

Martín del Campo, S. T., D. Picque, R. Cosío-Ramírez, and G. Corrieu. 2007. Evaluation of chemical parameters in soft mold-ripened cheese during ripening by mid-infrared spectroscopy. J. Dairy Sci. 90:3018–3027.

McQueen, D. H., R. Wilson, A. Kinnunen, and E. P. Jensen. 1995. Comparison of two infrared spectroscopic methods for cheese analysis. Talanta 42:2007–2015.

Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829.

Osborne, B. G. 2006. Near-infrared spectroscopy in food analysis. Pages 1–14 in Encyclopedia of Analytical Chemistry. John Wiley Sons Ltd., Hoboken, NJ.

Pérez, P., and G. de los Campos. 2014. Genome-wide regression and prediction with the BGLR statistical package. Genetics 198:483–495. https://doi.org/10.1534/genetics.114.164442.

Pi, F., H. Shinzawa, Y. Ozaki, and D. Han. 2009. Non-destructive determination of components in processed cheese slice wrapped with a polyethylene film using near-infrared spectroscopy and chemometrics. Int. Dairy J. 19:624–629.

Pierce, M. M., and R. L. Wehling. 1994. Comparison of sample handling and data treatment methods for determining moisture and fat in Cheddar cheese by near infrared spectroscopy. J. Agric. Food Chem. 42:2830–2835.

R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Accessed Dec. 2, 2015. http://www.R-project.org/.

Rodriguez-Otero, J. L., M. Hermida, and A. Cepeda. 1995. Determination of fat, protein, and total solids in cheese by near-infrared reflectance spectroscopy. J. AOAC Int. 78:802–806.

Sørensen, L. K., and R. Jepsen. 1998. Assessment of sensory properties of cheese by near-infrared spectroscopy. Int. Dairy J. 8:863–871.

Williams, P. 2003. Near-Infrared Technology—Getting the Best Out of Light: A Short Course in the Practical Implementation of Near Infrared Spectroscopy for the User. Ed. 1.1. PDK Projects Inc., Nanaimo, Canada.

Woodcock, T., C. C. Fagan, C. P. O'Donnell, and G. Downey. 2008. Application of near and mid infrared spectroscopy to determine cheese quality and authenticity. Food Bioprocess Technol. 1:117–129.

APPENDIX

Chemometric Procedure

Phenotypes were regressed on standardized spectra covariates using the following linear model:

$$y_i = \beta_0 + \sum_{j=1}^{n} x_{ij}\beta_j + \varepsilon_i,$$

where y_i represents the measured chemical and physical traits of the ith sample cheese, β_0 is an intercept, $\{x_{ij}\}$ are standardized spectra-derived wavelength data ($j = 1,\ldots, n$), β_j are the effects of each of the wavelengths, and ε_i are model residuals assumed to be *iid* (independent and identically distributed) with normal distribution centered at zero with variance σ_ε^2. Given the above assumption, the conditional distribution of the data given effects and variance parameters is

$$P(\mathbf{y}|\boldsymbol{\theta}) = \prod_{i=1}^{n} N\left(\mu_i, \sigma_\varepsilon^2\right),$$

where $\boldsymbol{\theta}$ represents the collection of model parameters $\boldsymbol{\theta} = \left\{\beta_0, \boldsymbol{\theta}, \sigma_\varepsilon^2\right\}$, $N\left(\mu_i, \sigma_\varepsilon^2\right)$ is a normal distribution centered at $\mu_i = \beta_0 + \sum_{j=1}^{n} x_{ij}\beta_j$ and with variance σ_ε^2, and $\boldsymbol{\beta} = \{\beta_j\}$ is a vector containing the effects of the individual spectra-derived wavelengths. Specification of the Bayesian model is completed by assigning prior distribution to the unknowns, $\boldsymbol{\theta}$. In the Bayesian models considered here, the prior density was as follows:

$$p\left(\boldsymbol{\theta}\right) = N\left(\beta_0|0, 1 \times 10^5\right)\chi^{-2}\left(\sigma_\varepsilon^2|\mathrm{df}_\varepsilon, S_\varepsilon\right)\left[\prod_{j=1}^{n}p\left(\beta_j|\Omega\right)\right]p\left(\Omega\right).$$

Here, the intercept is assigned a normal prior with a very large variance, the residual variance is assigned a scaled-inverse chi-squared (χ^{-2}) density with degree of freedom df_ε and scale parameters S_ε, and the effects of wavelengths are assigned *iid* priors, $p(\beta_j|\Omega)$, indexed by a set of hyper-parameters, Ω, which are also treated as random. The prior distribution assigned to the hyper-parameters is represented by $p(\Omega)$; $p(\beta_j|\Omega)$ and $p(\Omega)$ are different depending on the model implemented. In this work, the model used was Bayes B developed in the BGLR R package (de los Campos and Pérez-Rodriguez, 2014), where $p(\beta_j|\Omega)$ is a mixture of a point of mass at zero and a scaled t density, that is, $\left(\beta_j \mid \Omega\right) \stackrel{iid}{-} \pi \times t\left(\beta_j|\mathrm{df}_\beta, S_\beta\right) + \left(1-\pi\right) \times 1\left(\beta_j = 0\right)$; therefore, a priori, with probability π, β_j is drawn from the t density and with probability $(1 - \pi)\beta_j = 0$. We set $\mathrm{df}_\beta = 5$, and the other hyper-parameters were treated as random, specifically, $S_\beta \sim \mathrm{Gamma}(S_\beta|\mathrm{rate},\mathrm{shape})$ and $\pi \sim \mathrm{Beta}(\pi|\mathrm{shape}_1,\mathrm{shape}_2)$. All of these parameters were specified using built-in BGLR rules that select default values for these unknowns (θ) and are fully explained in Pérez and de los Campos (2014).

Scripts for Implementing Predictive Equations for the Bayes B Model in BGLR

This supplemental material contains instructions to obtain data and fit the models presented in the main manuscript. Appendix Tables A1, A2, A3, and A4 contain R scripts (R Core Team, 2013) for the Bayesian calibration using the R-package BGLR (de los Campos and Pérez-Rodriguez, 2014).

Table A1. Setting variables and parameters and partitioning data in training (TRN) and testing (TST) sets

Item	R script[1]
1	load("DataSpectra.RData")
2	# Assign the name of the y variable to the "trait" object, for example: Protein.
3	trait = "Protein"
4	# Define the number of iteration (nIter), number of samples to be discarded (burnIn)
5	# model (prior density) to be used: Bayes B
6	nIter = 10000
7	burnIn = 2000
8	model = "BayesB"
9	# Define the percentage of samples to be used as testing set
10	tst_perc = 0.20
11	y < −Y[,which(colnames(Y) = = trait)]
12	nTst < −round((length(y))*tst_percento,digits = 0)
13	n < −length(y)
14	tst < −sample(n, size = nTst, replace = FALSE)
15	# Creation of training and testing sets for the trait (Y) and for the spectra (X)
16	yTST < −y[tst]
17	XTST < −X[tst,]
18	yTRN < −y[−tst]
19	XTRN < −X[−tst,]

[1]R scripts (R Core Team, 2013) for the Bayesian calibration using the R-package Bayesian generalized linear regression (de los Campos and Pérez-Rodriguez, 2014).

Supplemental Data

The file DataSpectra.RData contains a portion of the data used in the scientific article and can be obtained by requesting it from the corresponding author.

The file contains the following R-objects:

- **Y**: Numeric vector (each datum is a sample) that contains the traits to be used for the calibrations. Each column represents a trait, and each row a cheese sample.
- **X**: Numeric matrix of spectra data where each row contains the spectrum for a cheese sample, the phenotype of which is located in the same row of the **Y** vector. Matrix **X** contains centered and standardized records. The detection of outliers is done using the Mahalanobis distances, and samples with a distance greater than 3 times the SD were discarded.

The identifiers of **Y** and **X** can be obtained with colnames(**Y**) (phenotype), rownames(**Y**) observation number, colnames(**X**) wavelength name.

Setting Up the Data and Model Fitting

Appendix Table A1 provides some initial values needed to use the BGLR package and partition the data in training and testing sets. We describe an example with a short number of iterations for illustration purposes only; however, the number of iterations and samples discarded as burn in should be determined by inspection of the trace plots of samples from the posterior.

Appendix Table A2 illustrates how to fit a simple model. For illustration purposes we present a very simple example; however, the Bayesian Generalized Linear Regression R-library (BGLR, available at CRAN; http://CRAN.R-project.org/package=BGLR) is a comprehensive statistical package that implements a large collection of Bayesian procedures. Here we are fitting a Bayes B model.

In Appendix Table A3 we illustrate how to calculate retrieve estimates from the model.

In Appendix Table A4 we illustrate how to calculate the predictive ability of the model in the testing set.

Table A2. Fitting the model

Item	R script[1]
1	library('BGLR')
2	ETA < −list(list(X = XTRN, model = model))
3	fm < −BGLR(y = yTRN, ETA = ETA, nIter = nIter, burnIn = burnIn, thin = 2)

[1]R scripts (R Core Team, 2013) for the Bayesian calibration using the R-package Bayesian generalized linear regression (BGLR; de los Campos and Pérez-Rodriguez, 2014).

STOCCO ET AL.

Table A3. Retrieving estimates

Item	R script[1]
1	# Extracting and plotting samples of the residual variance
2	fm$varE # residual variance
3	vare $< -$ scan('varE.dat')
4	plot(vare)
5	#Extracting effects associated to the wavelengths
6	wlef $< -$ fm$ETA[[1]]$b #coefficients
7	plot(wlef)

[1]R scripts (R Core Team, 2013) for the Bayesian calibration using the R-package Bayesian generalized linear regression (de los Campos and Pérez-Rodriguez, 2014).

Table A4. Evaluating prediction accuracy of the model in a testing set

Item	R script[1]
1	# Prediction of y on testing set applying the developed calibration equation
2	yHatTST $< -$XTST%*% fm$ETA[[1]]$b + fm$mu
3	#Descriptive statistics (mean and SD) for the training and the testing sets
4	meanyTRN $< -$ mean(yTRN)
5	sdTRN $< -$sd(yTRN)
6	meanyTST $< -$mean(yTST)
7	sdTST $< -$sd(yTST)
8	#Square correlation and error of validation
9	R2val $< -$(cor(yTST, yHatTST))^2
10	rmse_val $< -$sqrt((sum((yTST-yHatTST)^2))/length(yTST))

[1]R scripts (R Core Team, 2013) for the Bayesian calibration using the R-package Bayesian generalized linear regression (de los Campos and Pérez-Rodriguez, 2014).