



Near-infrared hyperspectral image analysis for monitoring the cheese-ripening process

Masoumeh Alinaghi,^{1,2} David Nilsson,¹ Nikita Singh,¹ Annika Höjer,³ Karin Hallin Saedén,³ and Johan Trygg^{1,4*}

¹Chemometrics Lab, Computational Life Science Cluster (CLiC), Umeå University, Umeå SE-901 87, Sweden

²Functional Microbiology, Institute of Microbiology, Department of Pathobiology, University of Veterinary Medicine, Vienna 1210, Austria

³Norrmejerier, Mejerivägen 2, Umeå SE-906 22, Sweden

⁴Sartorius Corporate Research, Sartorius, Sartorius Stedim Data Analytics, Umeå SE-903 33, Sweden

ABSTRACT

Ripening is the most crucial process step in cheese manufacturing and constitutes multiple biochemical alterations that describe the final cheese quality and its perceived sensory attributes. The assessment of the cheese-ripening process is challenging and requires the effective analysis of a multitude of biochemical changes occurring during the process. This study monitored the biochemical and sensory attribute changes of paraffin wax-covered long-ripening hard cheeses ($n = 79$) during ripening by collecting samples at different stages of ripening. Near-infrared hyperspectral (NIR-HS) imaging, together with free amino acid, chemical composition, and sensory attributes, was studied to monitor the biochemical changes during the ripening process. Orthogonal projection-based multivariate calibration methods were used to characterize ripening-related and orthogonal components as well as the distribution map of chemical components. The results approve the NIR-HS imaging as a rapid tool for monitoring cheese maturity during ripening. Moreover, the pixelwise evaluation of images shows the homogeneity of cheese maturation at different stages of ripening. Among the chemical compositions, fat content and moisture are the most important variables correlating to NIR-HS images during the ripening process.

Key words: near-infrared hyperspectral imaging, sensory analysis, cheese maturation, free amino acids, homogeneity distribution

INTRODUCTION

The maturation of long-ripening cheeses is expensive and time consuming, and its complicated biochemical process is difficult to control or predict (Fox et al.,

1996; Priyashantha et al., 2021b,c; Sun et al., 2022). During the ripening process, the curd turns into cheese with a particular flavor, taste, and texture based on the milk quality, microflora, processing steps, and storage conditions (Fox et al., 1996; Robinson and Wilbey, 1998; Rehn et al., 2010). In the maturation process, different physical, microbiological, and biochemical alternations, which are connected to the pH variation, lipid degradation, protein breakdowns, and accumulation of amino acids are taking place. The final flavor, taste, and texture of long-ripening cheeses are described by the end products of lipolysis, proteolysis, and glycolysis in the matured cheese, with proteolysis being the most important process (Fox, 1989). Cheese proteolysis involves the degradation of caseins, causing an increase in peptides and free amino acids during the ripening period (McSweeney and Fox, 1997). Glycolysis is mainly associated with lactose degradation (Bezerra et al., 2017) and lipolysis induces specific changes in fatty acid profile (Poveda and Cabezas, 2006; Park et al., 2007). The occurrence of all these biochemical processes is the main reason for the exclusive characteristics of different cheese varieties and indicates the usefulness of chemical and free amino acid analysis for evaluating cheese ripening at the molecular level.

Ripening is one of the critical processing steps that has an effect on the consistency and quality of the final product in cheese manufacture. Cheese characteristics, ripeness, and readiness for the market have been mainly evaluated by sensory panelists using conventional methods and subjective assessment of organoleptic properties (O'Shea et al., 1996; Coker et al., 2005). Sensory evaluation of the product can be considered a destructive, time-consuming, and expensive approach. Process analytical technology (PAT; Balboni, 2003) as a system for designing, analyzing, and controlling manufacturing by real-time measurements of critical quality and performance attributes has the ability to ensure the final product quality. Implementation of an online PAT system in the ripening stage of cheese production

Received February 15, 2023.

Accepted May 20, 2023.

*Corresponding author: johan.trygg@umu.se

can assist in achieving a consistently high-quality product. The near-infrared (NIR) hyperspectral (NIR-HS) imaging by employing data analytical methods has recently been used for the assessment of chemical composition and physical features of the sample in the area of food analysis (Gowen et al., 2007). The NIR-HS imaging is a rapid, nondestructive, low-cost, and chemical-free analysis that combines spectral and spatial information, which makes it ideal as a PAT tool.

In a 3-dimensional NIR-HS image, each pixel corresponds to spatialized spectral information combining chemical composition information with the analyte spatial distribution (Dorrepaaal et al., 2016). The NIR-HS imaging has been previously used for the prediction of macronutrients such as carbohydrates, fat, and proteins as well as the moisture content of products (Soto-Barajas et al., 2013; Currò et al., 2017). Predictive models on the NIR-HS images of long-ripening cheeses indicated the ability of NIR-HS image analysis to provide information on the alterations taking place during ripening, potentially allowing the end-date prediction of cheese maturation (Priyashantha et al., 2020, 2021a).

Advances in hyperspectral imaging techniques present remarkable opportunities for the analysis and characterization of spatial and spectral information in different types of samples (Gowen et al., 2007). However, the implementation of multivariate analysis is a necessity for extracting the whole spatial and structural information, especially in the complex matrix of food (Amigo et al., 2008; Amigo, 2010). Methods such as principal component analysis (PCA; Jolliffe, 2005; Grahn and Geladi, 2007), partial least squares (PLS; Wold et al., 2001), orthogonal projections to latent structures (OPLS; Trygg and Wold, 2002), and multivariate curve resolution–alternating least squares (MCR-ALS; de Juan and Tauler, 2006; Chang, 2007) enable the visualization of multiple biochemical constituents and their distribution in a complex surface. All these multivariate data analysis models are based on the analysis of combined spectral patterns from many wavelengths simultaneously, which has the advantage of dimension reduction and handling a high degree of noise.

Due to the incidence of different chemical processes during cheese ripening, understanding the evolution of chemical profiles and identifying the most remarkable changes throughout the ripening process would be crucial. There is potential for adaptation of PAT processing in the ripening stage of cheese production to improve the environmental sustainability and efficiency of the cheese-ripening process. The aim of this study is the development of a prediction model from hyperspectral imaging (HSI) to be applied online for monitoring cheese ripening to be able to evaluate maturity without

the destruction of cheese. In this paper, orthogonal projection-based multivariate calibration methods were used to develop and evaluate predictive models on the NIR-HS imaging data and the maturity values by trained sensory panelists for monitoring the ripening of long-ripening cheeses. Other parameters such as free amino acids and chemical compositions together with other sensory attributes such as taste, flavor, texture, and mouth feel were also evaluated during the ripening process.

MATERIALS AND METHODS

No human or animal subjects were used, so this analysis did not require approval by an Institutional Animal Care and Use Committee or Institutional Review Board.

Cheese Material and Experimental Design

This study was performed in the cheese manufacturing process at Norrmejerier, Sweden. Each cheese comprised an 18-kg cylinder (42 cm in diameter, 16 cm in height) which was brine-salted to a content of around 1.2% NaCl, coated with paraffin wax, and ripened under conditions explained by Rehn et al. (2010). Various cheese batches from different stages of ripening at 7 to 25 mo were evaluated for maturity quality by trained sensory panelists. However, the same cheese batches were not evaluated several times during ripening and therefore time-series data are not available. Six sensory panelists from the dairy company assessed each cheese against a standardized protocol reflecting on toughness, hardness, mouth feeling, crystals, bitterness, balance, aftertaste, and finally, overall cheese maturity on a scale from 1 to 9. The trained sensory panelists had a high level of sensitivity and consistency in the evaluation of the cheese sensory attributes. The performance of the panels was evaluated by standard deviations. The NIR-HS images and the chemical composition of cheeses such as fat content, protein, moisture, and pH were captured in parallel with the sensorial evaluation of the cheeses.

The procedure resulted in 79 images obtained from individual cheeses varying in ripeness (Supplemental Table S1, https://figshare.com/articles/online_resource/Near-infrared_hyperspectral_NIR-HS_image_analysis_for_monitoring_the_cheese_ripening_process/Supplementary_docx/24080802; Alinaghi, 2023). The free amino acid was analyzed by Société Générale de Surveillance SA through a so-called method of Amtliche Untersuchungsverfahren nach § 64 LFGB documented by the Federal Office of Consumer Protection and Food Safety (<https://www.bvl>

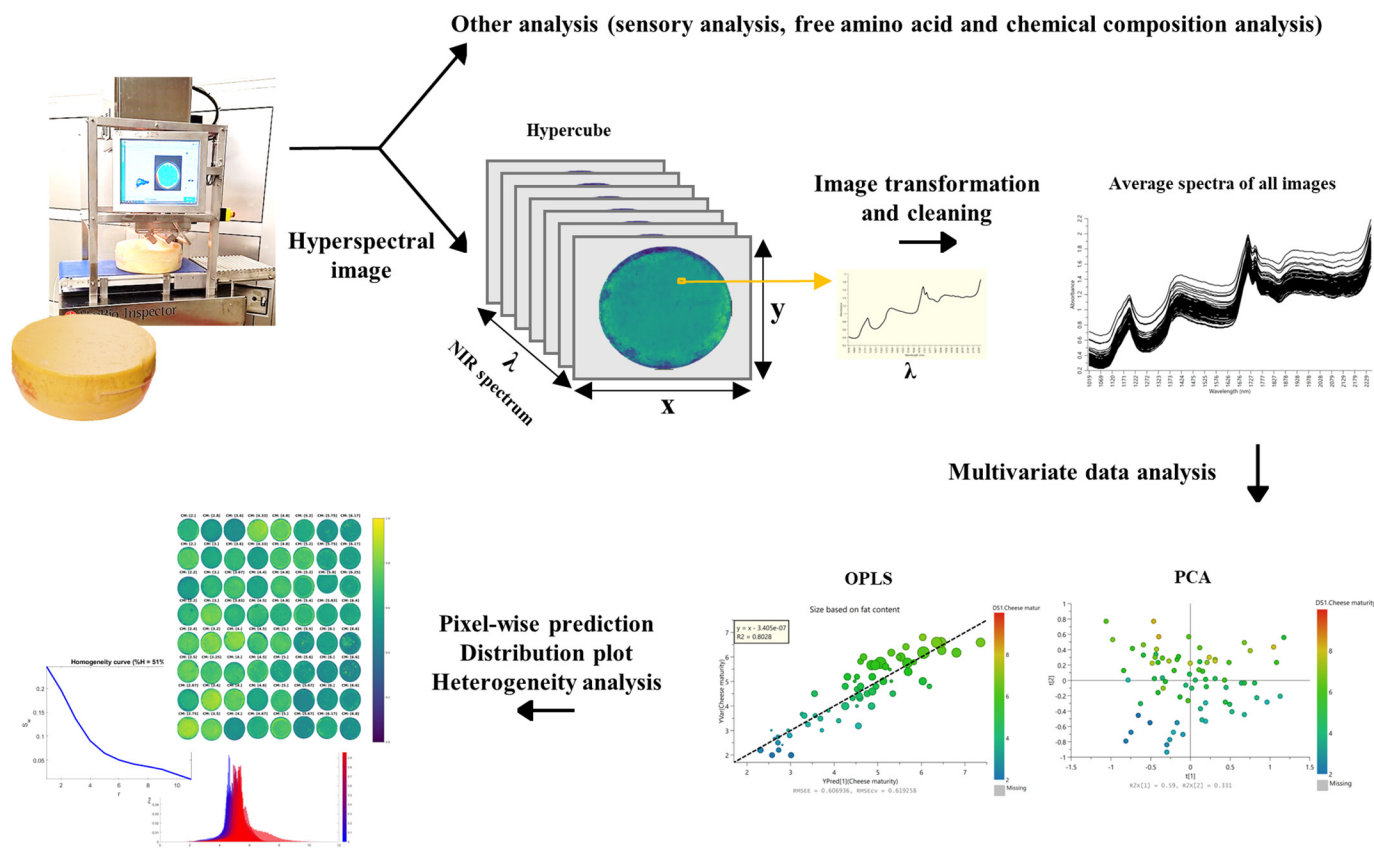


Figure 1. Illustration of hyperspectral imaging, hypercube data structure comprising wavelength (λ) and spatial (x and y) dimensions, and the model development. The near infrared hyperspectral (NIR-HS) images with some other techniques can be used to monitor the cheese-ripening process. The image planes at a few wavelengths as well as the spectrum of 1 selected pixel are schematically shown here. The average spectra of the image pixels are used for further multivariate data analysis techniques such as principal component analysis (PCA) and orthogonal projections to latent structures (OPLS) models as well as the heterogeneity analysis.

[.bund.de/DE/Arbeitsbereiche/01_Lebensmittel/01_Aufgaben/02_AmtlicheLebensmittelueberwachung/08_AmtlicheSammlung/lm_AmtlicheSammlung_node.html](https://www.bund.de/DE/Arbeitsbereiche/01_Lebensmittel/01_Aufgaben/02_AmtlicheLebensmittelueberwachung/08_AmtlicheSammlung/lm_AmtlicheSammlung_node.html)).

Hyperspectral Imaging Acquisition

An Umbio Inspector device (Umbio AB, Umeå, Sweden) was used for acquiring NIR-HS images (Hetta et al., 2017). All NIR-HS images were taken from cheeses covered with a 1-mm layer of paraffin wax (Priyashantha et al., 2020). The spectral range recorded was 937 to 2,542 nm, resulting in a NIR-HS image (378 pixels length \times 320 pixels width) in 256 wavelength channels. The scanning speed was set to acquire square pixels. Each image had roughly 121,000 pixels, of which approximately 63% were cheese pixels and 37% represented background pixels. A schematic illustration of the hyperspectral imaging process and hypercube data structure are represented in Figure 1.

Image Transformation and Cleaning

Reflectance images (I_{raw}) were recorded using the dark (I_{dark}) and white (I_{white}) reference data and were transformed into absorbance (A) by using Equation 1 (Grahn and Geladi, 2007):

$$A = -\log_{10} [(I_{\text{raw}} - I_{\text{dark}}) / (I_{\text{white}} - I_{\text{dark}})]. \quad [1]$$

The NIR-HS imaging captured a square-shaped image. Therefore, the area surrounding the circular cheeses was considered background information, giving rise to a noisy spectrum that required to be eliminated before any further analysis. The background pixels were eliminated by removing reflectance over 1.5 at band 55 (1,279 nm), to provide the highest difference between samples and the background (Priyashantha et al., 2021a). The average spectra of the image pixels were calculated for further analysis using in-house MatLab Script (version R2020b, MathWorks Inc.). Noisy wave-

lengths were removed by excluding wavelengths below 1,018 nm and above 2,250 nm. Therefore, 196 spectral wavelengths were included for further analysis.

Multivariate Data Analysis

Multivariate data analysis techniques such as PCA and OPLS were used in different parts of this study (Figure 1). Principal component analysis can reduce multivariate data to a smaller number of uncorrelated variables and provide an overview of the data. Although PLS regression represents a projection method to model the relationship between the \mathbf{X} data and \mathbf{Y} response, OPLS is an extension of the PLS method, which divides the systematic variation in the \mathbf{X} data matrix into 2 parts to provide an easier interpretation. One part models the correlation between \mathbf{X} and \mathbf{Y} , while the other part expresses the systematic variation in \mathbf{X} orthogonal to \mathbf{Y} .

To validate OPLS models against overfitting, the performance of the models was evaluated by the 7-fold cross-validation parameters Q^2 (goodness of prediction), R^2X , and R^2Y (goodness of fit). Multivariate data analysis such as PCA and OPLS was carried out using SIMCA (Sartorius Stedim Data Analytics, vs. 16.0.2, Umeå, Sweden) software. Other analyses including data visualization, unfolding, and reassembly after modeling were performed using MatLab (version R2020b, MathWorks Inc.).

The OPLS regression was used to study the relationship between NIR-HS images and the maturity of paraffin wax-covered, long-ripened hard cheeses. Both PCA and OPLS models, together with the D-optimal design (Mitchell, 2000), were performed to choose a selected number of samples for free amino acid analysis. An OPLS model was also built between the NIR-HS images and all sensory parameters, free amino acids of selected samples, and chemical composition as \mathbf{Y} variables to evaluate the importance of different variables in the prediction model.

Because the maturity criteria, sensory parameters, free amino acids, and chemical composition values were only available for whole cheeses, the images were replaced by average NIR spectra for regression analysis. The spectra were used in the standard normal variate-corrected and mean-centered form. The PCA was first performed on the average spectra to identify variations among the cheese samples.

The maturity values, as well as the orthogonal component to the cheese maturity, were predicted for all pixels using the OPLS calibration model with cheese maturity as y-variable. The distribution maps and histograms (i.e., distribution plots) of the predicted values were developed for each image. In addition, PCA was

performed on the histograms to identify variations among image distributions. The histogram data were aligned by shifting histogram plots using the interval-correlation-shifting algorithm (Savorani et al., 2010) and also mean-centered before PCA.

RESULTS AND DISCUSSION

PCA Model on Average Spectra

The PCA was performed on the average spectral data ($n = 77$, after removing 2 outliers) calculated from each image to identify variations among the samples (Figure 2). The PCA results revealed that the average images are changing by cheese maturity along the second principal component, with $t[1]$ and $t[2]$ explaining 59 and 33% variance of the average spectral data, respectively. This showed that a larger percentage of data variance (59%) is not correlated with the cheese maturity evolution. The third principal component explains 4% of the variation in the data and its loading plot does not show a meaningful spectral profile. Therefore, it can be concluded that the first 2 principal components explain the structured variation of the data. Elucidation of the first loading plot showed the **NIR** peaks related to the potentially wax content of the cheese (Priyashantha et al., 2021a). However, overlapping of the peaks in wax peaks with protein structure could be expected in this region (i.e., wavelengths 1,645–1,815 nm).

OPLS Model on Average Spectra

An OPLS model was developed between the average spectral data (\mathbf{X} , $n = 64$ NIR-HS images) and cheese maturity evaluated by sensory panelists (\mathbf{y}) (Figure 3). In OPLS analysis, 13 images were evaluated as overripened and therefore removed from further analysis. The overripened samples do not show a linear relationship with cheese maturity (Supplemental Figure S1, https://figshare.com/articles/online_resource/Near-infrared_hyperspectral_NIR-HS_image_analysis_for_monitoring_the_cheese_ripening_process/Supplementary_docx/24080802; Alinaghi, 2023). Nonlinear models of feedforward neural networks and support vector machines were also evaluated (but not shown) and no improvements were observed in the modeling results. Cross-validation for the OPLS model indicated the significance of one predictive and 2 orthogonal components corresponding to the $Q^2 = 0.78$ and root mean square error from cross-validation ($RMSE_{CV}$) = 0.62. The plot of predicted versus measured cheese maturity in Figure 3A and B illustrates that there was a strong correlation between the HSI data and the sensory maturity values. There-

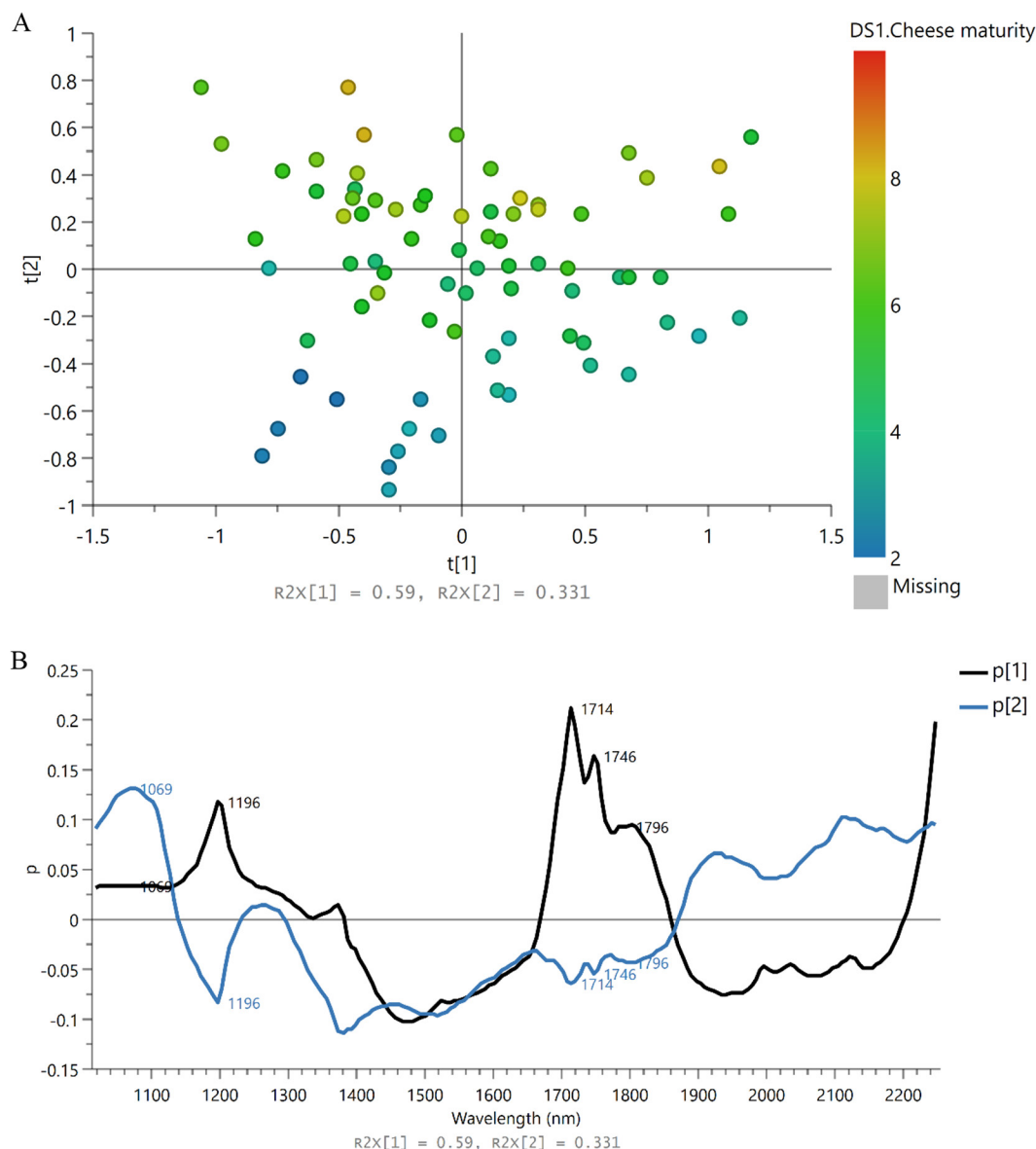


Figure 2. Principal component analysis score and loading plots of the average spectral data of cheese images by near-infrared hyperspectral imaging. Score plot colored according to cheese maturity (A). First ($p[1]$) and second ($p[2]$) loading plots (B). The $t[1]$ and $t[2]$ explain 59% and 33% variance of the average spectral data, respectively.

fore, NIR-HS image data can be used for the prediction of the maturity in paraffin wax-covered, long-ripening hard cheeses during the ripening process in a cheese manufacturing setting. However, the sensory analysis still needed to be performed before final approval to detect any sensory defects. The size of data points in Figure 3A and B is based on the fat content and moisture, respectively, illustrating the correlation of fat content and moisture with cheese maturity. This shows that cheese ripening causes an increase in fat content and a decrease in moisture over the process time. The plots of predicted versus measured cheese maturity

by protein and pH as data points' size are illustrated in Supplemental Figure S2 (https://figshare.com/articles/online_resource/Near-infrared_hyperspectral_NIR-HS_image_analysis_for_monitoring_the_cheese_ripening_process_Supplementary_docx/24080802;Alinaghi, 2023), which shows that the changes in protein and pH are not correlated with cheese maturity. The OPLS predictive component, which shows the correlation of the maturity variable to the data explains 29% of the data variation, causing 71% of the data variation unexplained by the maturity. Thus, the orthogonal components could be inspected to understand

the unexplained variation in the data. Figure 3C and D illustrate the orthogonal score plot (i.e., $to[2]$ vs. $to[1]$) and are colored by protein and pH values, respectively, suggesting that the orthogonal variation in the data can also be explained by factors other than the wax content of the cheese (Figure 3E).

Further regression model between the NIR-HS data and all sensory parameters, free amino acids, and chemical composition was also considered. However, as there was not possibility to perform the free amino acid analysis for all samples, a selection of samples was chosen. Samples were selected based on the variation observed in the NIR-HS images. Therefore, all images were averaged into mean spectra and an OPLS model ($R^2X = 0.72$, $R^2Y = 0.98$, 4 predictive + 1 orthogonal component) was constructed between the mean spectra and 6 key response parameters (cheese maturity, fat, moisture, protein, DM, and pH). In addition, a 3-component PCA model ($R^2X = 0.96$) was generated. The scores of 4 predictive components of the OPLS model, together with the 3 scores of the PCA model, were assembled as a candidate set, or 7 factors in total. A D-optimal design (Mitchell, 2000), using the G-efficiency criterion, was used for selecting 44 samples for further amino acid analysis.

Finally, a 3-component OPLS model (i.e., one predictive and 2 orthogonal components) between the average spectra (X , $n = 64$ NIR-HS images) and all sensory parameters, free amino acids for selected samples, and chemical composition as Y variables (30 Y variables) was also built (Figure 4). The cross-validation resulted in the significance of one predictive component, indicating a high level of correlation between the variables as well as $Q^2 = 0.58$ and $RMSE_{CV} = 0.78$. The X/Y overview plot (Figure 4A) displays the individual R^2 and Q^2 for every y -variable with an OPLS model. All sensory attributes and most of the free amino acids have quite similar results. Among the chemical composition variables, the fat content and afterward moisture have the highest Q^2 and R^2 values. The protein content does not show good results, and its calculated standard error is also higher. Even though the variation in fat content might be originated from the batch differences, a high Q^2 value represents the correlation of fat content to NIR-HS image data. Finally, the results of Figure 4A highlight the potential of NIR-HS imaging for monitoring the cheese composition during maturation. The plot in Figure 4B also shows that the fat content is higher in cheeses with a higher level of maturity, whereas the moisture and protein are showing decreased value. The pH of cheeses remained unchanged during ripening. For the free amino acids, all except glutamic acid show an increasing pattern with an enhanced level of maturity. This could be explained due the degradation of proteins

to free amino acids (McSweeney and Fox, 1997), whereas the conversion of glutamic acid to glutamate leads to its decrease during the maturation process. The performance of this OPLS model (Figure 4) is comparable with the OPLS model with only cheese maturity as a y -variable (Figure 3), as the model resulted in similar Q^2 and R^2 values for cheese maturity. However, the model with cheese maturity as a y -variable has better potential for implementation in the industrial setting for online monitoring due to less number of variables needed for the calibration model and thus simplicity of the model.

Image Visualization and Distribution Map

Even though analyzing the average images could provide valuable information about the maturity process, the full potential of the 3-dimensional data is not revealed yet. Therefore, comparing component-wise image distributions at different cheese maturity stages might be informative. Quantitative estimates of different components' concentration in each pixel were predicted by using the OPLS model with cheese maturity as a y -variable from section 3.2. To provide comparability, the images were normalized to the maximum value of each cheese. Figure 5A visualizes the maturity distribution maps of all individual cheeses sorted by the cheese maturity from low to high level.

As NIR-HS image analysis can provide physicochemical and spatial information about samples, NIR-HS imaging is an ideal analytical technique for performing heterogeneity analysis. Heterogeneity, described as the variability of predicted component across the image space, were studied by histograms to observe the scatter around a central value of image pixels. Histograms provide information on the intrinsic (constitutional) variability of the pixel measurements in the sample (Petersen et al., 2005; de Moura França et al., 2017). Histograms of the predicted image pixels (Figures 5B, D, and E) suggest that cheese maturation is homogeneous compared with the histogram of the orthogonal distribution (i.e., the maturity-related distribution map within each cheese has less scattering in concentration values within an image). The maturity-related distribution maps also reveal a shift in the mean with the maturity level, which is also illustrated by PCA score plot of the histograms (Figures 5B and C). However, the orthogonal component shows nonhomogeneous constitutional distribution (Figures 5E). These results illustrate that the nonhomogeneity in the original images can be explained by the components orthogonal to the cheese maturity. Conversely, distributional heterogeneity (de Moura França et al., 2017), which can be studied by a homogeneity curve, evaluates how uniformly

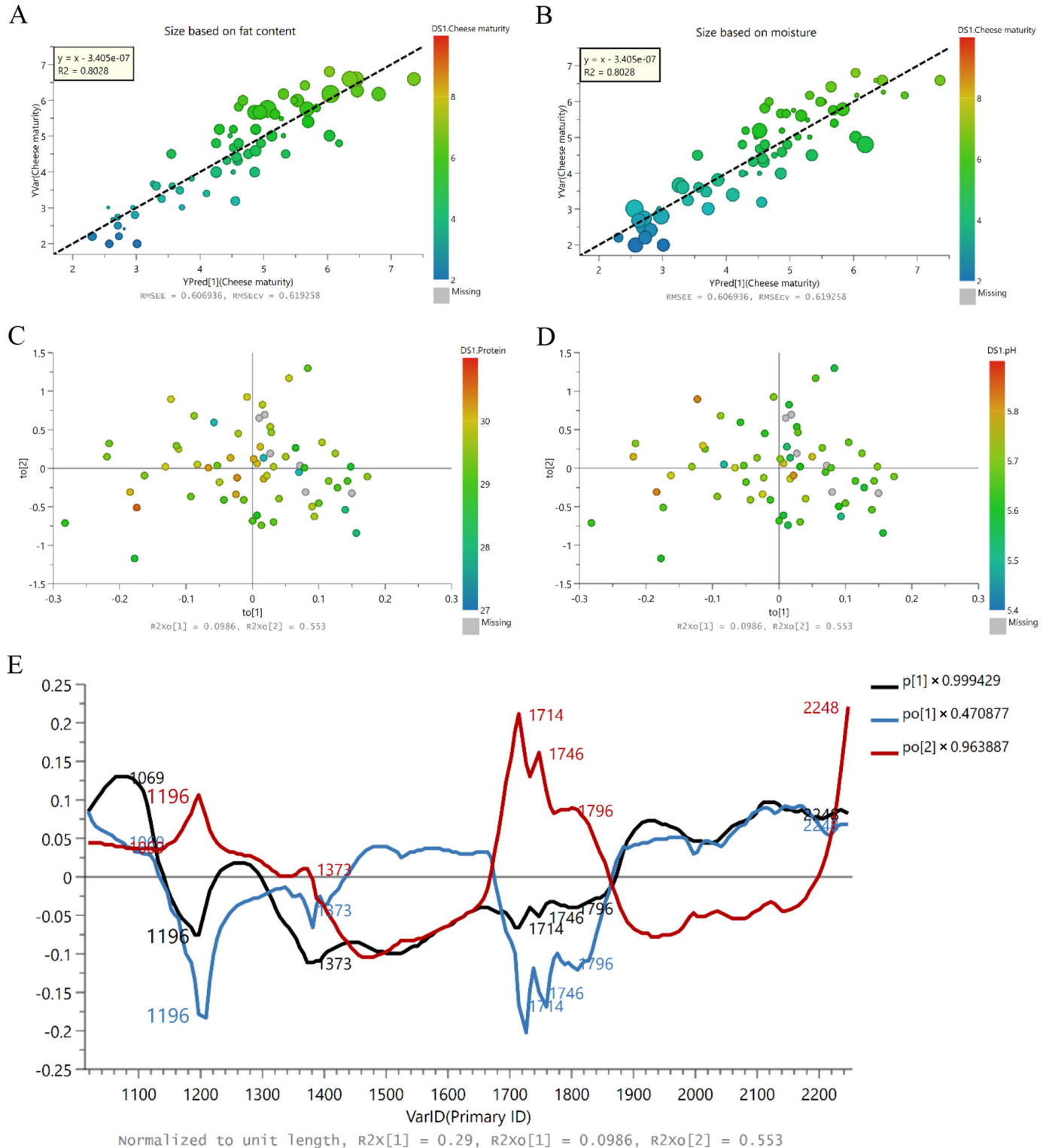


Figure 3. Orthogonal projections to latent structures (OPLS) model with 3 components on the average spectra (X , $n = 64$ near-infrared hyperspectral images after removing the overripened samples) with 196 spectral wavelengths and cheese maturity as Y variable. The OPLS model parameters are as follows: R^2X for the predictive component = 0.29, R^2X for the first orthogonal component = 0.1, R^2X for the second orthogonal component = 0.55, R^2 for the predictive component = 0.8, and $Q^2 = 0.78$. Predicted versus measured cheese maturity colored by the cheese maturity and size by fat content (A), predicted versus measured cheese maturity colored by the cheese maturity and size by moisture (B), OPLS score plots of $to[2]$ versus $to[1]$, colored based on the protein content (C), OPLS score plots of $to[2]$ versus $to[1]$, colored based on the pH (D), and OPLS loading of the predictive and orthogonal components (E).

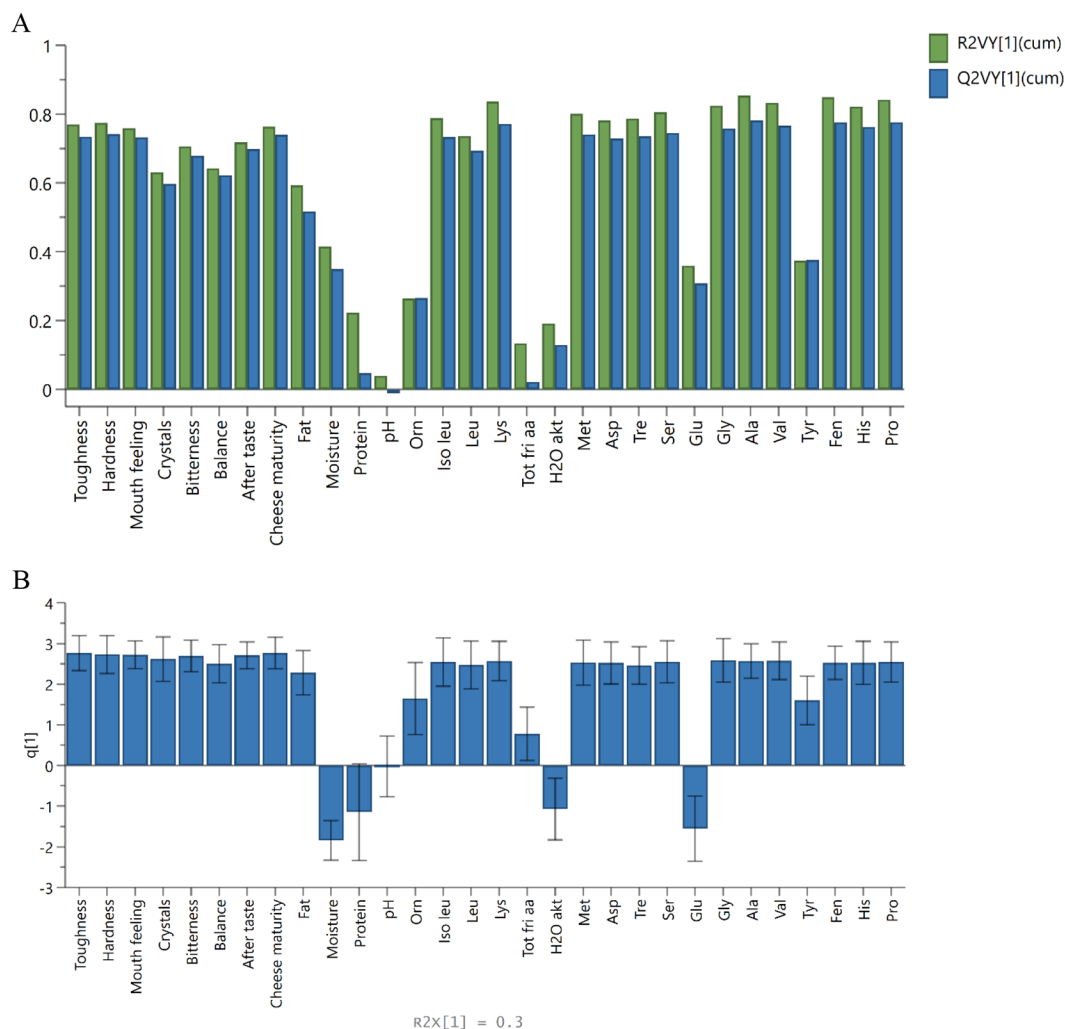


Figure 4. Orthogonal projections to latent structures (OPLS) model with one predictive and 2 orthogonal components on the average spectra (\mathbf{X} , $n = 58$ NIR-HS images after removing the overripened samples and the samples with missing \mathbf{Y} variables) with 196 spectral wavelengths and all sensory parameters, free amino acids, and chemical composition (30 \mathbf{Y} variables). The OPLS model parameters are as follows: R^2X for predictive component = 0.3, R^2X for first orthogonal component = 0.54, R^2X for second orthogonal component = 0.11, R^2 for predictive component = 0.59, $Q^2 = 0.58$. The \mathbf{X}/\mathbf{Y} overview plot presents R^2 and Q^2 values for different \mathbf{Y} variables (A), and \mathbf{Y} matrix loading plot (B).

the different constituents are distributed in the image by taking the properties of the pixel neighborhood into account. Homogeneity curves (Figures 5D and E) show a higher level of distributional homogeneity in the maturity-related distribution maps compared with the distribution maps of the orthogonal component. Thus, heterogeneity analysis might be a valuable tool for defining the quality of end products, describing the evolution of processes and detecting possible abnormal behaviors. However, application of heterogeneity analysis for better decision-making and quality control in monitoring cheese ripening in the industrial setting needs further investigations.

Metabolic processes and the progress of the biochemical reactions are mainly responsible for the flavor

and texture change during the maturation, indicating the effect of the storage time. However, the ripening rate of the cheese might be different even by excluding the maturation condition and variation in the cheese composition (Muir et al., 1995). Therefore, modeling the HSI data with the time as a y-variable could cause different results, due to differences in the maturation rate. Skeie et al. (2006) reported different modeling results for free amino acid prediction when the age of the cheese was used as a predictor compared with NIR. Priyashantha et al. (2020) have previously shown a PLS model between the cheese age and the HSI data, considering that the maturity index and cheese age show a linear relationship until 18 mo (~550 d). Analyzing time-series data should be with careful consideration

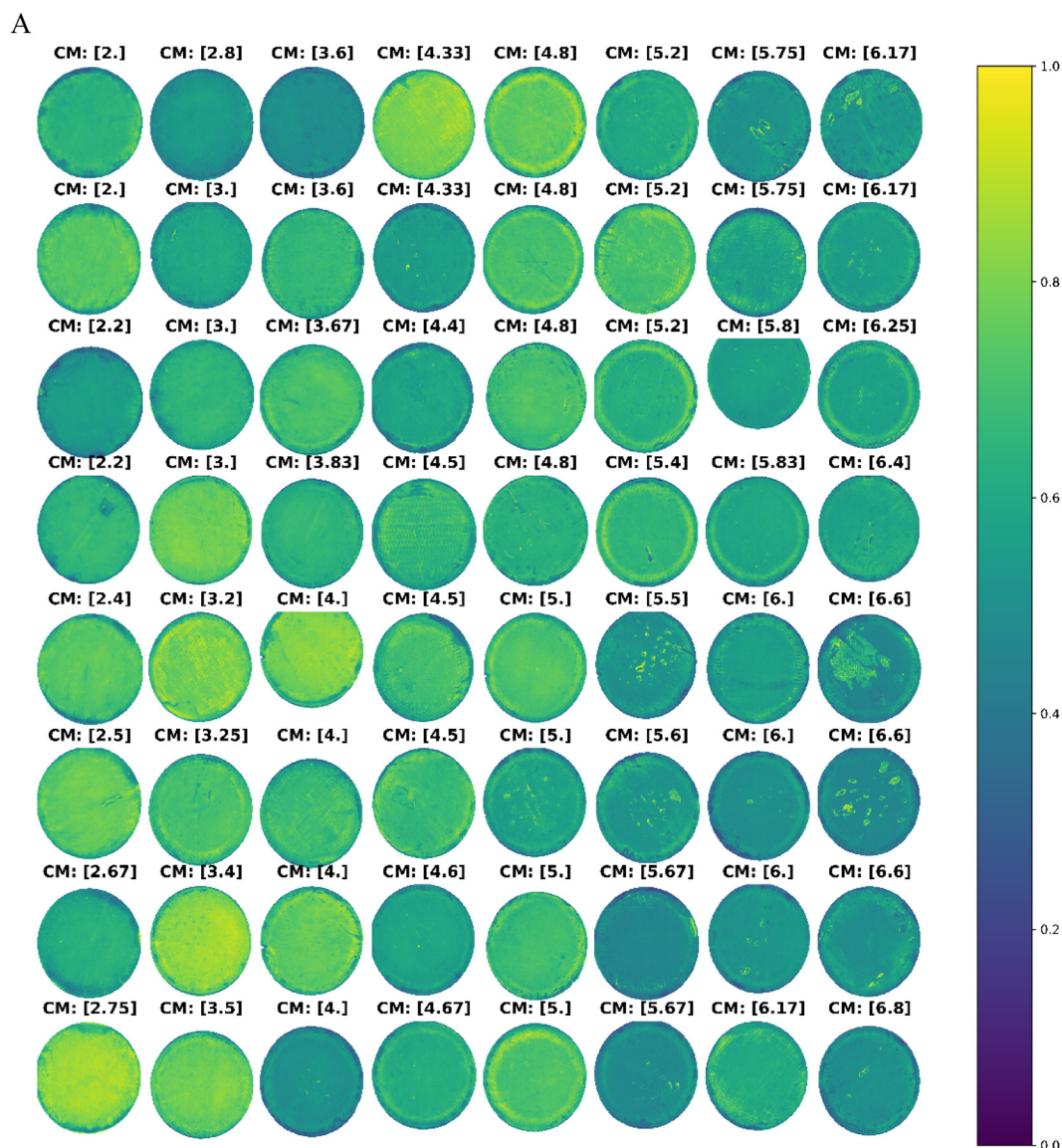


Figure 5. Heterogeneity information obtained from pixelwise prediction of the orthogonal projections to latent structures (OPLS) model with cheese maturity as Y variable. The distribution maps of the predicted component are labeled by cheese maturity (CM) values (A). Constitutional heterogeneity is represented by histograms obtained from map concentration values, colored according to the cheese maturity value with a color range of blue for low maturity and red for high maturity (B), the results of principal component analysis on the histograms, $t[1]$ and $t[2]$ explain 58% and 28% variance of the data (C), Distribution maps, histograms, and homogeneity curves for the maturity-related (D) and orthogonal component (E) of a selected image with cheese maturity value of 2.0 (top left image in panel A).

as some degree of the correlation between time-related changes and maturity-related changes might be observed (Alinaghi et al., 2022). However, the time-series data evaluating the same cheese batches over maturation time is not available in this study. Therefore, analysis of the present data set does not struggle with the time effect because the cheese maturity values from the sensory panelist were used as the y-variable. This experimental design has also the benefit that the drift in the sensory evaluation sessions is small due to the

avoidance of the time effect. Therefore, evaluation of the attribute and variables describing the maturity stages of cheese is possible. However, the time-series analysis and tracking of the chemical changes of the cheese over maturation time cannot be feasible in this experimental setting.

For the evaluation of the prediction models, the importance of a panel's performance in the sensory data cannot also be underestimated. Lack of consistency among the panelists can cause a higher value

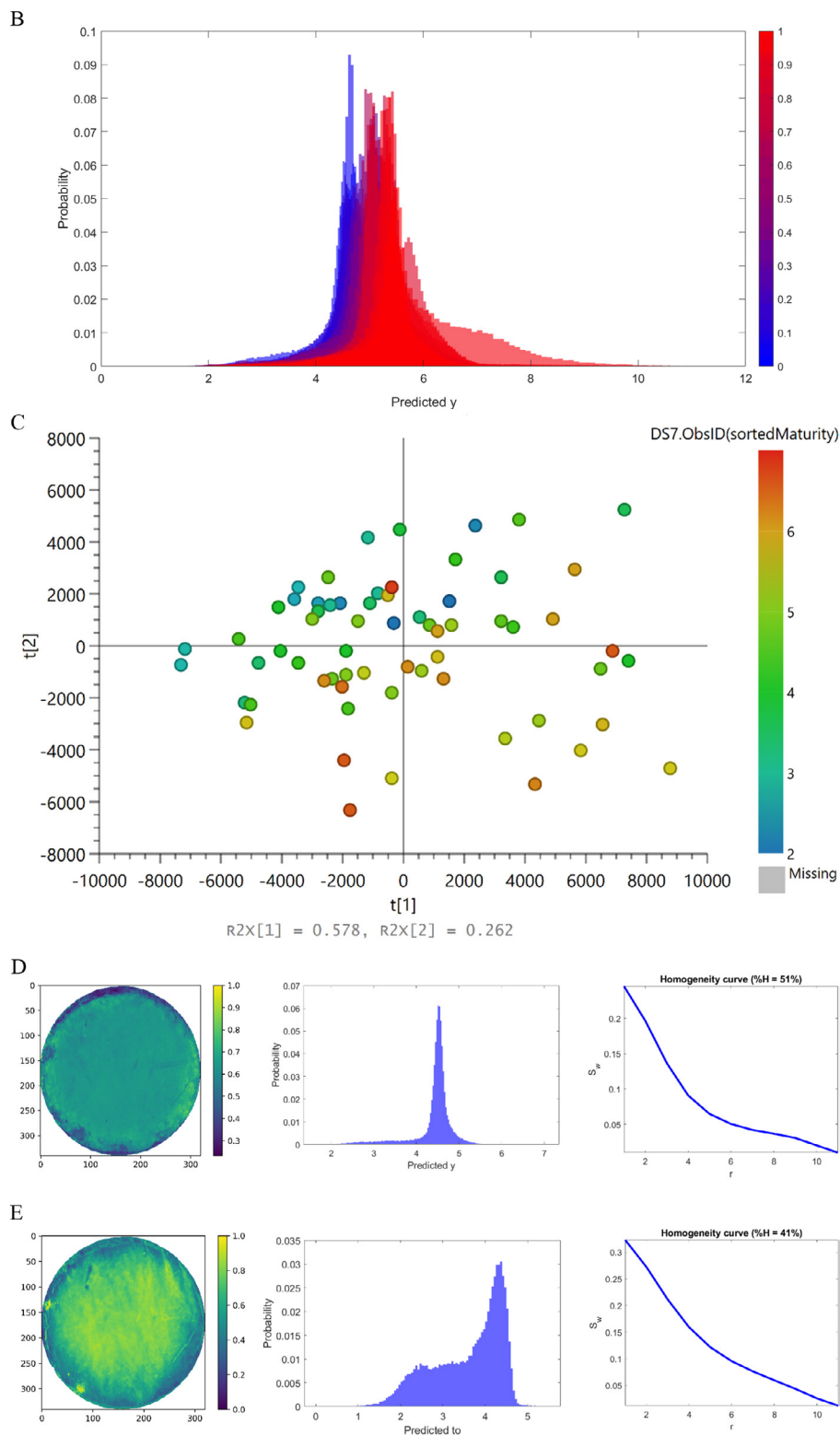


Figure 5 (Continued). Heterogeneity information obtained from pixelwise prediction of the orthogonal projections to latent structures (OPLS) model with cheese maturity as Y variable. The distribution maps of the predicted component are labeled by cheese maturity (CM) values (A). Constitutional heterogeneity is represented by histograms obtained from map concentration values, colored according to the cheese maturity value with a color range of blue for low maturity and red for high maturity (B), the results of principal component analysis on the histograms, $t[1]$ and $t[2]$ explain 58% and 28% variance of the data (C), Distribution maps, histograms, and homogeneity curves for the maturity-related (D) and orthogonal component (E) of a selected image with cheese maturity value of 2.0 (top left image in panel A).

of RMSE_{CV} in the modeling and difficulty in the interpretation of data. The reproducibility as the ability to score cheeses samples averagely the same as other panelists can be used as a measure for outlier detection, detecting patterns in the measures as well as the overall panel's performance (Rossi, 2001). The performance can be tracked for different attributes over time to determine whether a panelist can perform properly. The performance of the sensory panelist was evaluated by calculating the standard deviation of panel scores. The results (Supplemental Figure S3, https://figshare.com/articles/online_resource/Near-infrared_hyperspectral_NIR-HS_image_analysis_for_monitoring_the_cheese_ripening_process_Supplementary_docx/24080802; Alinaghi, 2023) showed that the panel had a similar standard deviation throughout all sessions, and no outlier was detected. However, individual differences in sensory perceptions will remain even after through training of panelists. Therefore, a level of uncertainty can be expected in the OPLS modeling results causing prediction errors in the model.

CONCLUSIONS

Monitoring cheese ripening is challenging and requires analysis of multiple biochemical alterations occurring during the process. Cheese ripening has been mainly assessed by conventional methods of sensory analysis which are destructive, time consuming, and expensive. However, final flavor, taste, and texture of long-ripening cheeses will be characterized by end products of biochemical processes. The analysis results highlight the potential of NIR-HS image analysis for monitoring the cheese composition and free amino acids during maturation. Therefore, NIR-HS imaging together with data analytical methods can be employed online as a rapid tool for monitoring cheese maturity. Moreover, the pixelwise evaluation of NIR-HS images can reveal the homogeneity of cheese maturation at different stages of ripening. This improved understanding of undergoing biochemical processes will facilitate further adoption of monitoring approaches and PAT tools in cheese manufacture. However, this leads to the question that how heterogeneity analysis may provide a tool for better decision-making and quality control in cheese production.

ACKNOWLEDGMENTS





This study received no external funding. The authors have not stated any conflicts of interest.

REFERENCES

- Alinaghi, M. 2023. Near-infrared hyperspectral (NIR-HS) image analysis for monitoring the cheese ripening process_Supplementary.docx. figshare. Online resource. <https://doi.org/10.6084/m9.figshare.24080802.v1>.
- Alinaghi, M., I. Surowiec, S. Scholze, C. McCready, C. Zehe, E. Johansson, J. Trygg, and O. Cloarec. 2022. Hierarchical time-series analysis of dynamic bioprocess systems. *Biotechnol. J.* 17:2200237 <https://doi.org/10.1002/biot.202200237>.
- Amigo, J. M. 2010. Practical issues of hyperspectral imaging analysis of solid dosage forms. *Anal. Bioanal. Chem.* 398:93–109. <https://doi.org/10.1007/s00216-010-3828-z>.
- Amigo, J. M., J. Cruz, M. Bautista, S. Maspocho, J. Coello, and M. Blanco. 2008. Study of pharmaceutical samples by NIR chemical-image and multivariate analysis. *Trends Analyt. Chem.* 27:696–713. <https://doi.org/10.1016/j.trac.2008.05.010>.
- Balboni, M. L. 2003. Process analytical technology. *Pharm. Technol.* 27:54–67.
- Bezerra, T. K. A., N. M. de Oliveira Arcanjo, E. F. Garcia, A. M. P. Gomes, R. de Cássia Ramos do Egypto Queiroga, E. L. de Souza, and M. S. Madruga. 2017. Effect of supplementation with probiotic lactic acid bacteria, separately or combined, on acid and sugar production in goat 'coalho'cheese. *Lebensm. Wiss. Technol.* 75:710–718. <https://doi.org/10.1016/j.lwt.2016.10.023>.
- Chang, C.-I. 2007. *Hyperspectral Data Exploitation: Theory and Applications*. John Wiley & Sons.
- Coker, C., R. Crawford, K. Johnston, H. Singh, and L. Creamer. 2005. Towards the classification of cheese variety and maturity on the basis of statistical analysis of proteolysis data—a review. *Int. Dairy J.* 15:631–643. <https://doi.org/10.1016/j.idairyj.2004.10.011>.
- Curro, S., C. L. Manuelian, M. Penasa, M. Cassandro, and M. De Marchi. 2017. Technical note: Feasibility of near infrared transmittance spectroscopy to predict cheese ripeness. *J. Dairy Sci.* 100:8759–8763. <https://doi.org/10.3168/jds.2017.13001>.
- de Juan, A., and R. Tauler. 2006. Multivariate curve resolution (MCR) from 2000: progress in concepts and applications. *Crit. Rev. Anal. Chem.* 36:163–176. <https://doi.org/10.1080/10408340600970005>.
- de Moura França, L., J. M. Amigo, C. Cairós, M. Bautista, and M. F. Pimentel. 2017. Evaluation and assessment of homogeneity in images. Part 1: Unique homogeneity percentage for binary images. *Chemom. Intell. Lab. Syst.* 171:26–39. <https://doi.org/10.1016/j.chemolab.2017.10.002>.
- Dorrepaal, R., C. Malegori, and A. Gowen. 2016. Tutorial: Time series hyperspectral image analysis. *J. Near Infrared Spectrosc.* 24:89–107. <https://doi.org/10.1255/jnirs.1208>.
- Fox, P. 1989. Proteolysis during cheese manufacture and ripening. *J. Dairy Sci.* 72:1379–1400. [https://doi.org/10.3168/jds.S0022-0302\(89\)79246-8](https://doi.org/10.3168/jds.S0022-0302(89)79246-8).
- Fox, P. F., J. M. Wallace, S. Morgan, C. M. Lynch, E. J. Niland, and J. Tobin. 1996. Acceleration of cheese ripening. *Antonie van Leeuwenhoek* 70:271–297. <https://doi.org/10.1007/BF00395937>.
- Gowen, A. A., C. P. O'Donnell, P. J. Cullen, G. Downey, and J. M. Frias. 2007. Hyperspectral imaging—An emerging process analytical tool for food quality and safety control. *Trends Food Sci. Technol.* 18:590–598. <https://doi.org/10.1016/j.tifs.2007.06.001>.
- Grahn, H., and P. Geladi. 2007. Clustering and classification in multispectral imaging for quality inspection of postharvest products. Pages 43–67 in *Techniques and Applications of Hyperspectral Image Analysis*. John Wiley & Sons.
- Hetta, M., Z. Mussadiq, J. Wallsten, M. Halling, C. Swensson, and P. Geladi. 2017. Prediction of nutritive values, morphology and agronomic characteristics in forage maize using two applications of NIRS spectrometry. *Acta Agric. Scand. B Soil Plant Sci.* 67:326–333. <https://doi.org/10.1080/09064710.2017.1278782>.
- Jolliffe, I. 2005. Principal component analysis. Pages 542–543 in *Encyclopedia of Statistics in Behavioral Science*. B. S. Everitt and D. C. Howell, ed. John Wiley & Sons.

- McSweeney, P., and P. Fox. 1997. Chemical methods for the characterization of proteolysis in cheese during ripening. *Lait* 77:41–76. <https://doi.org/10.1051/lait:199713>.
- Mitchell, T. J. 2000. An algorithm for the construction of “D-optimal” experimental designs. *Technometrics* 42:48–54. <https://doi.org/10.1080/00401706.2000.10485978>.
- Muir, D., E. Hunter, J. Banks, and D. Horne. 1995. Sensory properties of Cheddar cheese: changes during maturation. *Food Res. Int.* 28:561–568. [https://doi.org/10.1016/0963-9969\(95\)00039-9](https://doi.org/10.1016/0963-9969(95)00039-9).
- O’Shea, B. A., T. Uniacke-Lowe, and P. Fox. 1996. Objective assessment of Cheddar cheese quality. *Int. Dairy J.* 6:1135–1147. [https://doi.org/10.1016/0958-6946\(95\)00065-8](https://doi.org/10.1016/0958-6946(95)00065-8).
- Park, Y., M. Juárez, M. Ramos, and G. Haenlein. 2007. Physico-chemical characteristics of goat and sheep milk. *Small Rumin. Res.* 68:88–113. <https://doi.org/10.1016/j.smallrumres.2006.09.013>.
- Petersen, L., P. Minkinen, and K. H. Esbensen. 2005. Representative sampling for reliable data analysis: Theory of sampling. *Chemom. Intell. Lab. Syst.* 77:261–277. <https://doi.org/10.1016/j.chemolab.2004.09.013>.
- Poveda, J., and L. Cabezas. 2006. Free fatty acid composition of regionally-produced Spanish goat cheese and relationship with sensory characteristics. *Food Chem.* 95:307–311. <https://doi.org/10.1016/j.foodchem.2004.12.045>.
- Priyashantha, H., A. Höjer, K. H. Saedén, Å. Lundh, M. Johansson, G. Bernes, P. Geladi, and M. Hetta. 2020. Use of near-infrared hyperspectral (NIR-HS) imaging to visualize and model the maturity of long-ripening hard cheeses. *J. Food Eng.* 264:109687. <https://doi.org/10.1016/j.jfoodeng.2019.109687>.
- Priyashantha, H., A. Höjer, K. H. Saedén, Å. Lundh, M. Johansson, G. Bernes, P. Geladi, and M. Hetta. 2021a. Determining the end-date of long-ripening cheese maturation using NIR hyperspectral image modelling: A feasibility study. *Food Control* 130:108316. <https://doi.org/10.1016/j.foodcont.2021.108316>.
- Priyashantha, H., Å. Lundh, A. Höjer, G. Bernes, D. Nilsson, M. Hetta, K. H. Saedén, A. H. Gustafsson, and M. Johansson. 2021b. Composition and properties of bovine milk: A study from dairy farms in northern Sweden; Part I. Effect of dairy farming system. *J. Dairy Sci.* 104:8582–8594. <https://doi.org/10.3168/jds.2020-19650>.
- Priyashantha, H., Å. Lundh, A. Höjer, G. Bernes, D. Nilsson, M. Hetta, K. H. Saedén, A. H. Gustafsson, and M. Johansson. 2021c. Composition and properties of bovine milk: A study from dairy farms in northern Sweden; Part II. Effect of monthly variation. *J. Dairy Sci.* 104:8595–8609. <https://doi.org/10.3168/jds.2020-19651>.
- Rehn, U., M. A. Petersen, K. H. Saedén, and Y. Ardö. 2010. Ripening of extra-hard cheese made with mesophilic DL-starter. *Int. Dairy J.* 20:844–851. <https://doi.org/10.1016/j.idairyj.2010.06.001>.
- Robinson, R., and R. Wilbey. 1998. Cheese maturation. Pages 193–231 in *Cheesemaking Practice*. R. K. Robinson and R. A. Wilbey. Springer.
- Rossi, F. 2001. Assessing sensory panelist performance using repeatability and reproducibility measures. *Food Qual. Prefer.* 12:467–479. [https://doi.org/10.1016/S0950-3293\(01\)00038-6](https://doi.org/10.1016/S0950-3293(01)00038-6).
- Savorani, F., G. Tomasi, and S. B. Engelsen. 2010. icoshift: A versatile tool for the rapid alignment of 1D NMR spectra. *J. Magn. Reson.* 202:190–202. <https://doi.org/10.1016/j.jmr.2009.11.012>.
- Skeie, S., G. Feten, T. Almøy, H. Østlie, and T. Isaksson. 2006. The use of near infrared spectroscopy to predict selected free amino acids during cheese ripening. *Int. Dairy J.* 16:236–242. <https://doi.org/10.1016/j.idairyj.2005.03.008>.
- Soto-Barajas, M. C., M. I. González-Martín, J. Salvador-Esteban, J. M. Hernández-Hierro, V. Moreno-Rodilla, A. M. Vivar-Quintana, I. Revilla, I. L. Ortega, R. Morón-Sancho, and B. Curto-Diego. 2013. Prediction of the type of milk and degree of ripening in cheeses by means of artificial neural networks with data concerning fatty acids and near infrared spectroscopy. *Talanta* 116:50–55. <https://doi.org/10.1016/j.talanta.2013.04.043>.
- Sun, L., Å. Lundh, A. Höjer, G. Bernes, D. Nilsson, M. Johansson, M. Hetta, A. H. Gustafsson, K. H. Saedén, and J. Dicksved. 2022. Milking system and premilking routines have a strong effect on the microbial community in bulk tank milk. *J. Dairy Sci.* 105:123–139. <https://doi.org/10.3168/jds.2021-20661>.
- Trygg, J., and S. Wold. 2002. Orthogonal projections to latent structures (O-PLS). *J. Chemometr.* 16:119–128.
- Wold, S., M. Sjöström, and L. Eriksson. 2001. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* 58:109–130. [https://doi.org/10.1016/S0169-7439\(01\)00155-1](https://doi.org/10.1016/S0169-7439(01)00155-1).

ORCIDS

Masoumeh Alinaghi  <https://orcid.org/0000-0001-9304-1239>
 David Nilsson  <https://orcid.org/0000-0003-3957-6499>
 Annika Höjer  <https://orcid.org/0000-0002-3526-2081>
 Johan Trygg  <https://orcid.org/0000-0001-6611-4607>