Current Research in Food Science 4 (2021) 18–27

Contents lists available at ScienceDirect

Current Research in Food Science

journal homepage: www.editorialmanager.com/crfs/

Research Paper

Principal component analysis of hyperspectral data for early detection of mould in cheeselets

Jessica Farrugia [a], Sholeem Griffin [a,b,c], Vasilis P. Valdramidis [b,c], Kenneth Camilleri [a,d], Owen Falzon [a,*]

[a] Centre for Biomedical Cybernetics, University of Malta, Msida, Malta
[b] Department of Food Sciences and Nutrition, University of Malta, Msida, Malta
[c] Centre of Molecular Medicine and Biobanking, University of Malta, Msida, Malta
[d] Department of Systems & Control Engineering, Faculty of Engineering, University of Malta, Msida, Malta

ARTICLE INFO

Keywords:
Hyperspectral imaging
Principal component analysis
Agar
Cheeselet

ABSTRACT

The application of non-destructive process analytical technologies in the area of food science got a lot of attention the past years. In this work we used hyperspectral imaging to detect mould on milk agar and cheese. Principal component analysis is applied to hyperspectral data to localise and visualise mycelia on the samples' surface. It is also shown that the PCA loadings obtained from a set of training samples can be applied to hyperspectral data from new test samples to detect the presence of mould on these. For both the agar and cheeselets, the first three principal components contained more than 99 % of the total variance. The spatial projection of the second principal component highlights the presence of mould on cheeselets. The proposed analysis methods can be adopted in industry to detect mould on cheeselets at an early stage and with further testing this application may also be extended to other food products.

1. Introduction

High quality food and food safety are important goals for any food industry. Assessing the quality of dairy products is a challenge that dairy industries face since the shelf life for such products is typically short. Conventionally, multiple chemical and microbiological contamination tests are carried out to assess the quality of dairy products. However, these assessment methods typically involve destructive testing and are time consuming (Burke et al., 2018)). Plate counting can also be used along microbial analysis of dairy products to monitor spoilage. This technique measures the level of microorganism in dairy products. However, a limitation in this method is that an undetectable amount of food contaminants can still be present (Parseelan et al., 2019). Other techniques such as the polymerase chain reaction (PCR) can also be used to identify bacterial pathogens in dairy products and they were found to be more sensitive and quicker when compared to microbial culture based techniques (McLean et al., 2010). However, these molecular techniques are quite expensive and highly trained scientists are required to implement such techniques.

Fast, non-invasive and non-destructive methods can make the assessment of dairy food product much more efficient. In particular, the use of hyperspectral imaging for contaminant detection in food products as this offers a quick, non expensive and non-destructive solution (Sun, 2010). Hyperspectral imaging combines spatial and spectroscopic information providing a spectral representation for each spatial location in the imaged area (Feng and Sun, 2012). Thus, hyperspectral imagers can provide a data cube (or hypercube) consisting of three-dimensional data, namely two spatial dimensions and one spectral dimension. When an object is exposed to electromagnetic radiation, the incident radiation is reflected, transmitted, absorbed, or scattered, depending on its chemical composition and physical structure. When used to image food, the light component that is reflected by the food and captured by the hyperspectral imager, can be used to assess the properties of the imaged food and the presence of contaminants and their spatial distribution on the food (Sun, 2010).

While various studies have considered the use of hyperspectral imaging for food contamination inspection (Barbin et al., 2013; Del Fiore et al., 2010; Feng et al., 2018; Feng and Sun, 2013; Foca et al., 2016; Huang et al., 2013; Jiang et al., 2016; Kamruzzaman et al., 2015; Tao et al., 2015; Wang et al., 2014, 2015), the application of the technology

* Corresponding author.
E-mail address: owen.falzon@um.edu.mt (O. Falzon).

https://doi.org/10.1016/j.crfs.2020.12.003
Received 10 September 2020; Received in revised form 27 December 2020; Accepted 31 December 2020

on dairy products such as cheese has been rather limited (Gowen et al, 2009, 2011; Lei and Sun, 2019). To date no other study involving the detection of fungal contamination in cheese using hyperspectral imaging has been reported. A few studies have been published making use of non-sterile model systems or commercial cheese for the assessment of adulteration with starch or transglutaminase, for the assessment of cheese ripening, or for cheese classification (Barreto et al., 2018; Calvini et al., 2020; Lei et al., 2019; Shan et al., 2020). The limited use of hyperspectral imaging with dairy products may be due to the limited adoption of this imaging modality within the field of food science. This modality requires further development of algorithms to detect aberrations from desirable characteristics, before application to dairy products.

This study investigated the use of hyperspectral imaging for the detection of contamination in cheeselets. *Penicillium chrysogenum*, a contaminant found in commercial cheese, was used as the food contaminant. *P. chrysogenum* or *P. notatum* (formerly) is commonly isolated from temperate and subtropical environments and can grow on salted food products (Samson et al., 2010). It is often isolated in indoor environments, in particularly damp or water-damaged buildings (Andersen et al., 2011). The environment in which certain artisanal cheeses are made often have high humidity levels. Therefore, *P. chrysogenum* lends itself to being a persistent contaminant of commercial cheese as its growth is conducive to a salty food matrix like cheese, which is aged in an indoor environment under humid conditions. Moreover, a recent study found that 27% of all hard and 17% of all soft sheeps' milk cheese to be contaminated with the *Penicillium* genera, which further validates this choice of food contaminant (Griffin et al., 2020). Apart from tests on actual cheeselets, a number of tests were carried out on milk agar samples. The milk agar was used as a controlled model system before assessment of cheeselets. The milk agar contained ampicillin, which stops the growth of other micro-organisms and serves as a surface for the *P. chrysogenum* to grow on as it would on cheese. The surface of the milk agar is smooth and uniform when compared to cheese, which makes the identification of contaminants easier. The milk agar experiment served as a proof of concept for the use of the hyperspectral imaging method to identify fungal contaminants on the samples, before proceeding to the more complex food matrix of cheeselets.

A range of methods have been used to analyse hyperspectral data to detect fungi on different food. These include principal component analysis (PCA), partial least squares regression (PLSR), non-linear regression method of back propagation neural network (BPNN), linear discriminant analysis (LDA), quadratic discriminant analysis (QDA)and factorial discriminant analysis (FDA) (Feng and Sun, 2013; Jiang et al., 2016; Del Fiore et al., 2010; MA et al., 2014).

In this study, PCA was used to analyse the hyperspectral imaging data obtained for the milk agar and cheeselets. Hyperspectral data is often analysed and processed using PCA in order to reduce the dimensionality of the data. The method allows for the extraction of a limited set of features that are representative of salient characteristics in the data, such as an optimal set of wavelengths (Del Fiore et al., 2010; MA et al., 2014; Feng et al., 2018; Jiang et al., 2016; Munir et al., 2018). In contrast with methods that adopt a supervised approach for model learning, PCA provides an unsupervised data-driven approach that thus does not require the use of any prior training data. In this work, PCA was applied to the hyperspectral data from the milk agar and cheeselets in order to spatially localise and visualise mycelia at the pixel level. The visualisation of the resulting principal components allowed for the early detection of contaminated regions on cheeselets.

2. Materials and methods

2.1. Sample preparation

In total 13 agar samples and 12 cheeselets were prepared for hyperspectral imaging and analysis. The purpose of this study was to challenge the PCA testing model. A known proportion of samples were contaminated with known contaminant. The end point result was a dichotomous one that is contaminated or uncontaminated. Thus, for this reason the number of samples used for this study was sufficient to achieve such aim. The details of the preparation methods for both the milk agar as well as the cheeselets is provided in the following sections.

2.1.1. Preparation of agar samples

A 500 mL Duran bottle containing 100 mL of a 15% (w/v) agar-agar (Scharlau, Spain) solution was sterilised by autoclaving at 121°C for 15 min. Raw sheep's milk was sterilised and aseptically added to the 100 mL of 15% (w/v) agar-agar solution to give a 1.5% (w/v) milk agar in which (0.5 mL of 100 mg/mL) ampicillin (HiMedia, India) stock solution was added. Milk agar discs were then made in 5 cm petri dishes. A sterility test was performed by storing three milk agar plates in a 25°C incubator for 5 days.

2.1.2. Preparation of cheeselets

A cheese model was developed using raw sheep's milk provided by a local farmer. The milk was used to make a 1% (w/v) NaCl solution, which was sterilised by heating to 78°C for 2 min, then rapidly cooled to 40°C by submersion into an ice bath. The sterilised milk:salt solution was inoculated with 0.05% (w/v) rennet and commercial Lactobacillus spp cultures and incubated for 1 h at ambient temperature, allowing the milk to form curds. The curds were transferred to sterile cheese moulds under aseptic conditions and allowed to drain in a laminar flow cabinet for 1 h. The cheese moulds containing the drained curds were submerged to a sterile 28% (w/v) NaCl salt bath aseptically and incubation for 2 min at ambient temperature. They were then removed from the salt bath aseptically and drained for a further 24 h in a laminar flow cabinet. The cheese was removed from the moulds and transferred aseptically to a sterile container and inoculated before further use.

2.1.3. Preparation of fungal stock

A contaminated commercial sample was used for the isolation of the fungal contaminate of Maltese cheese "ġbejna". The contaminated cheese was weighed aseptically in a BagMixer bag and four volumes of Ringer's solution were transferred to it for homogenisation in a BagMixer for 180 s.

The cheese homogenate was serially diluted in Ringer's solution, 100 μl inoculated on plated out on Sabouraud's dextrose agar (SDA) and incubated at 25°C for five days. A single fungal colony was isolated on SDA and used to generate a glycerol spore stock in a sterile 40% (v/v) glycerol solution, which was stored at −20°C until further use. The fungal contaminant isolated from the commercial cheese was later identified as *P. chrysogenum* by 18S rRNA gene sequencing.

2.1.4. Sample inoculation

The spore stock was quantified using a counting chamber and diluted to 10^6 CFU.mL^{-1} in 40% (v/v) glycerol solution to create a working spore stock. Milk agar plates were inoculated with 25 μl of the working spore stock, which was spread using a sterile spreader. The model cheese was inoculated with 10 μl of the same working spore stock in duplicate on the opposite sides of the cheeselets. The same stock was used for both agar and cheese. Different volumes of spore stock were used as to prevent the surface of the cheeselets to be completely covered with mould, to represent the type of contamination present on commercial cheese. Inoculated milk agar and model cheese were incubated at 25°C until required for hyperspectral analysis. Sterile 40% (v/v) glycerol solution was used as a negative control for contamination. Ten milk agar discs were inoculated with the working spore stock, whereas three discs were used for the negative control. Similarly, ten model cheese were inoculated with the working spore stock and two samples were used as a negative control. The samples were incubated at 25°C and imaged daily with an interval of 24 h.

2.2. Experimental setup and data acquisition

A SPECIM IQ portable hyperspectral camera covering a spectral band of 400–1000 nm with a spectral resolution of 7 nm was used for this study. The imaging of the samples was done in a dark room. Controlled lighting provided by four halogen lamps was used for a homogeneous illumination of the sample of interest and to avoid shadows or specular reflections.

Images of all milk agar and cheeselets were taken on consecutive days to monitor the progression of *P. chrysogenum* growth on samples. The size of the acquired raw hyperspectral datacube was of 512 × 512 × 204 (samples × lines × bands).

Hyperspectral images of the inoculated and control agar samples were acquired over a period of four consecutive days. Specifically a single image for each of the 13 agar samples (3 control and 10 inoculated samples) was acquired every 24 h, for a total of 52 images. On the second day of imaging, mycelia was present on the agar surface and the inoculated samples became sporulated on the third and fourth day of imaging.

Hyperspectral images of the inoculated and control cheeselets were acquired over a period of two consecutive days. A single image for each of the 12 cheeselets (2 control and 10 inoculated samples) was acquired every 24 h, for a total of 24 images. On the first day of imaging, mycelia was present on the cheeselets and these samples became sporulated on the second day of imaging.

2.3. Hyperspectral data analysis

2.3.1. Preprocessing

The dark current effect of the camera was removed by covering the lens of the camera, acquiring a dark image, and subtracting the hyperspectral data for the dark image from the hyperspectral data acquired from the target of interest. A white diffuse reflectance target was used to obtain a white reference image.

The relative reflectance (R) of the hyperspectral data was then obtained from:

$$R = \frac{I_s - I_d}{I_w - I_d} \times 100\% \quad (1)$$

where I_s is the raw hyperspectral data, I_d is the dark image and I_w is the white reference image.

An elliptical spatial region of interest (ROI) was selected from the hyperspectral data for futher processing. The ROI consisted of the whole sample's surface excluding the background and the samples' border. Fig. 1 shows a representation of this.

The hyperspectral data from the ROI was then further processed using PCA and the resulting components were evaluated to determine the spectral characteristics and spatial distribution of the agar and cheeselet contaminants. The details of the PCA implementation in this application are next presented.

2.3.2. Principal component analysis of hyperspectral data

PCA yields an orthogonal transformation that converts a set of observations from a set of potentially correlated variables into a set of linearly uncorrelated variables referred to as principal components (PC). The transformed PCs are ordered in such a way that the first one accounts for the highest variability in the data and each subsequent component contains the remaining amount of variance in descending order. All of the PCs are orthogonal to each other.

In this work, PCA was applied directly to the pixel data from the hyperspectral imaging process. As a preprocessing step the hyperspectral data cube was restructured as shown in Fig. 2. The pixels of the ROI can be considered as a set of correlated variables to which PCA was applied. The score matrix Z is given by:

$$Z = XW \quad (2)$$

where the rows of the input matrix $X \in \mathbb{R}^{K \times L}$ represent the spectral values for K = M×N pixels over L spectral bands. $W \in \mathbb{R}^{L \times P}$ is the loading matrix, the columns of which represent the eigenvectors of the covariance matrix of X. The columns of W provide the transformation functions that map the pixel spectral vectors into P PCs. The columns of $Z \in \mathbb{R}^{K \times P}$ represent the PC scores which are the representations of X in the PC space. Each PC image is the product between the pixel spectral vectors of X and a column of W. Each PC image is obtained by reshaping each PC making up Z, to a two dimensional representation.

PCA was applied directly to the hyperspectral data obtained from all agar and cheeselets (considering all control and contaminated samples). For each of these, the score images together with their corresponding spectra were computed to detect and identify mycelial growth and spores on the milk agar and cheeselets. Additionally, spectral signatures of the areas containing mycelial growth and uncontaminated areas for the ten contaminated cheeselets were computed and investigated. All processing and analysis was carried out using MATLAB.

2.3.3. Predetermined principal component loadings

Apart from applying PCA to each individual sample, an investigation was also carried out to determine whether the transformation coefficients obtained from one sample could effectively be applied to another sample for contamination detection. This investigation was meant to determine if the transformation coefficients obtained from a set of contaminated samples used for training can be used to identify the presence of mould in unseen test samples.

The PCA was first trained on five agar samples containing only mould at mycelial growth. The transformation coefficients for the first five PC scores were determined for each trained sample. The transformation

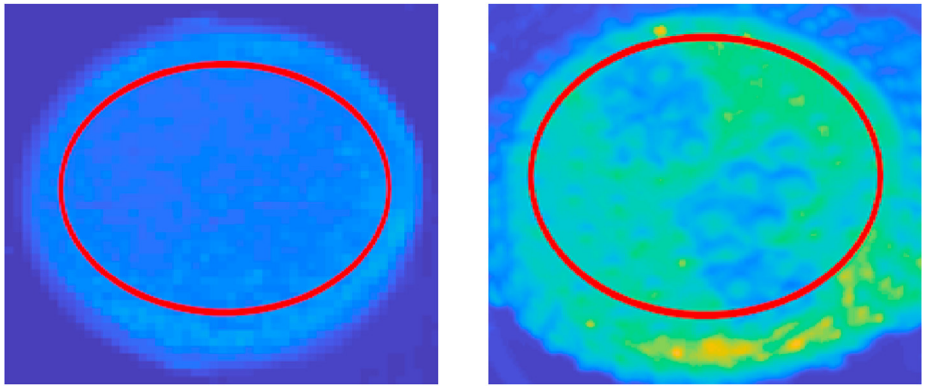

Fig. 1. An example of an agar sample (left) and a cheeselet (right) with the ROI (red ellipse) selected. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

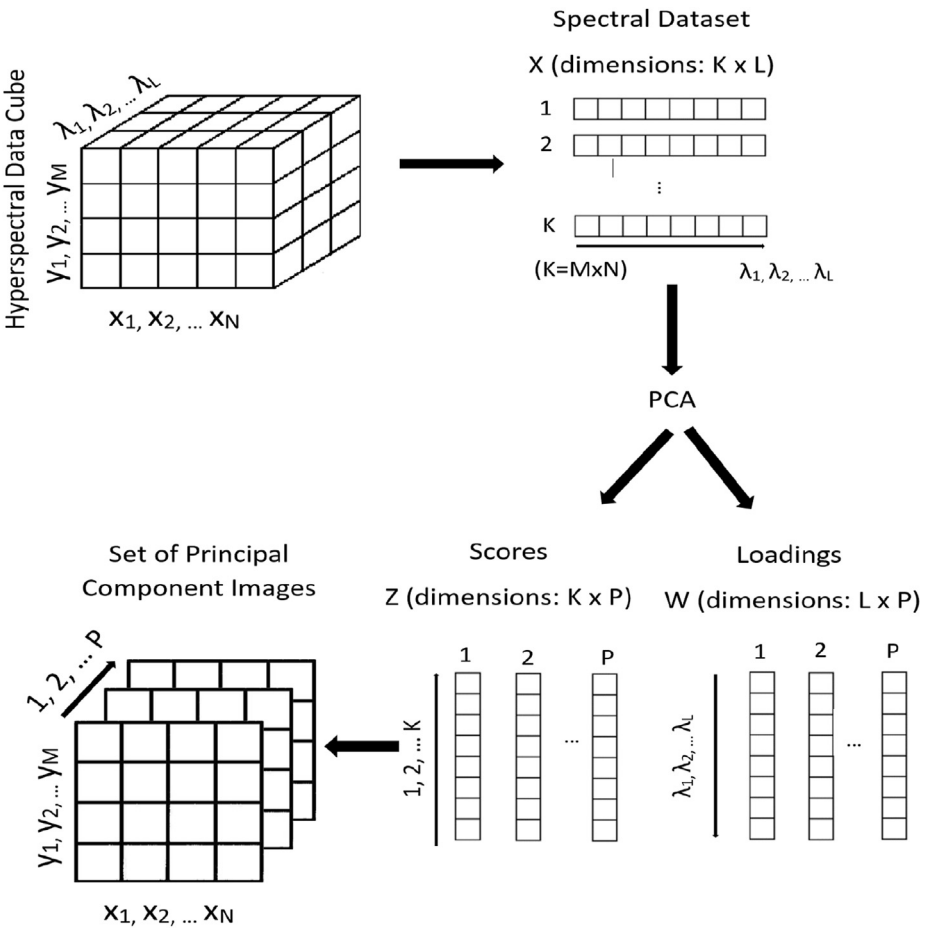

Fig. 2. Reshaping of the hyperspectral data cube as input for PCA. x_i represents the spatial column i = {1, 2, …, N}, y_j represents the spatial row j = {1, 2, …, M}, and λ_l represents the l^{th} spectral band where l = {1, 2, …, L}. The data cube is reshaped to form a K × L matrix where K = M × N. After PCA is applied to X, a score and a loading matrix are obtained. The score is given by a K × P matrix and loading is defined by an L × P matrix. P is the number of PCs. Score data can be reshaped to obtain a maximum of P PC images.

coefficients obtained from the samples used for training were also applied to the raw matrix $X \in \mathbb{R}^{K \times L}$ of other samples having the same type of mould as the trained samples. The resulting images from the PC scores were visually inspected and compared to the score images of the same samples as a result of the direct application of PCA on the raw data to assess whether the same contaminants were highlighted at the same spot. The trained sample which showed the most consistent spectral loadings as the other samples for the five PCs was selected. Its transformation coefficients were tested on new samples. The projected score images of the PC scores were visually inspected to determine if any regions highlighted in these images actually reflected the presence of sporulated areas which eventually showed up.

The transformation coefficients obtained from the samples used for training were also applied to the control samples to assess whether the presence of mould would show up in these cases.

The same procedures was repeated for the cheeselets. However, in this case the PCA was trained on ten cheeselets containing mycelia.

3. Results

3.1. PCA results on agar samples

The first three PCs represented more than 99% of the variance for the control agar samples, agar samples that contain mycelia (samples imaged on day 2) and agar samples that contain sporulated areas (samples imaged on day 4).

Fig. 3a shows an RGB representation of the agar samples considered, together with the first three PCs loadings (Fig. 3b, d and 3f) and the corresponding spatial projections (Fig. 3c, e and 3g) that result when applying PCA to the corresponding hyperspectral data. The agar sample imaged on day 2 contained mycelia whilst the same agar sample imaged on day 4 became sporulated. All 10 samples that contained mycelia showed a similar spectral representation across the whole spectrum in the first three PCs. The same can be said for the agar samples containing sporulated areas. When comparing the first PC (Fig. 3b) of the three different cases of agar, one can note that there are significant difference in the 400 nm–500 nm range and in the 900 nm–1000 nm range. The spectral representation of the second PC (Fig. 3d) for the control sample and the sample containing mycelia is similar in the 400 nm–900 nm range but shows distinct differences in the 900 nm–1000 nm range. There are amplitude value differences when comparing the second PC loadings of the sample containing sporulated areas to the control sample (Fig. 3d). The spectral representation of the third PC is different for the three different cases of agar.

The score images (Fig. 3c) associated with the first PC show the general illumination gradient effect of the agar samples including the presence of any shadows and bright spots. On the other hand, the score images associated with the second PC (Fig. 3e) highlight mouldy areas on the samples and it is more sensitive for mycelial growth. For the agar samples, the score images associated with third PC (Fig. 3g) highlight the texture of the mould on the samples. These characteristics in the score images were observed in all agar samples.

3.2. PCA results on cheeselets

The first three PCs for cheese samples (all contamination cases) also explained more than 99 % of the total variance. Fig. 4 presents an RGB representation (Fig. 4a) of the cheeselets considered, together with the first three PCs loadings (Fig. 4b, d and 4f) and their corresponding spatial projections (Fig. 4c, e and 4g) that result when applying PCA to the

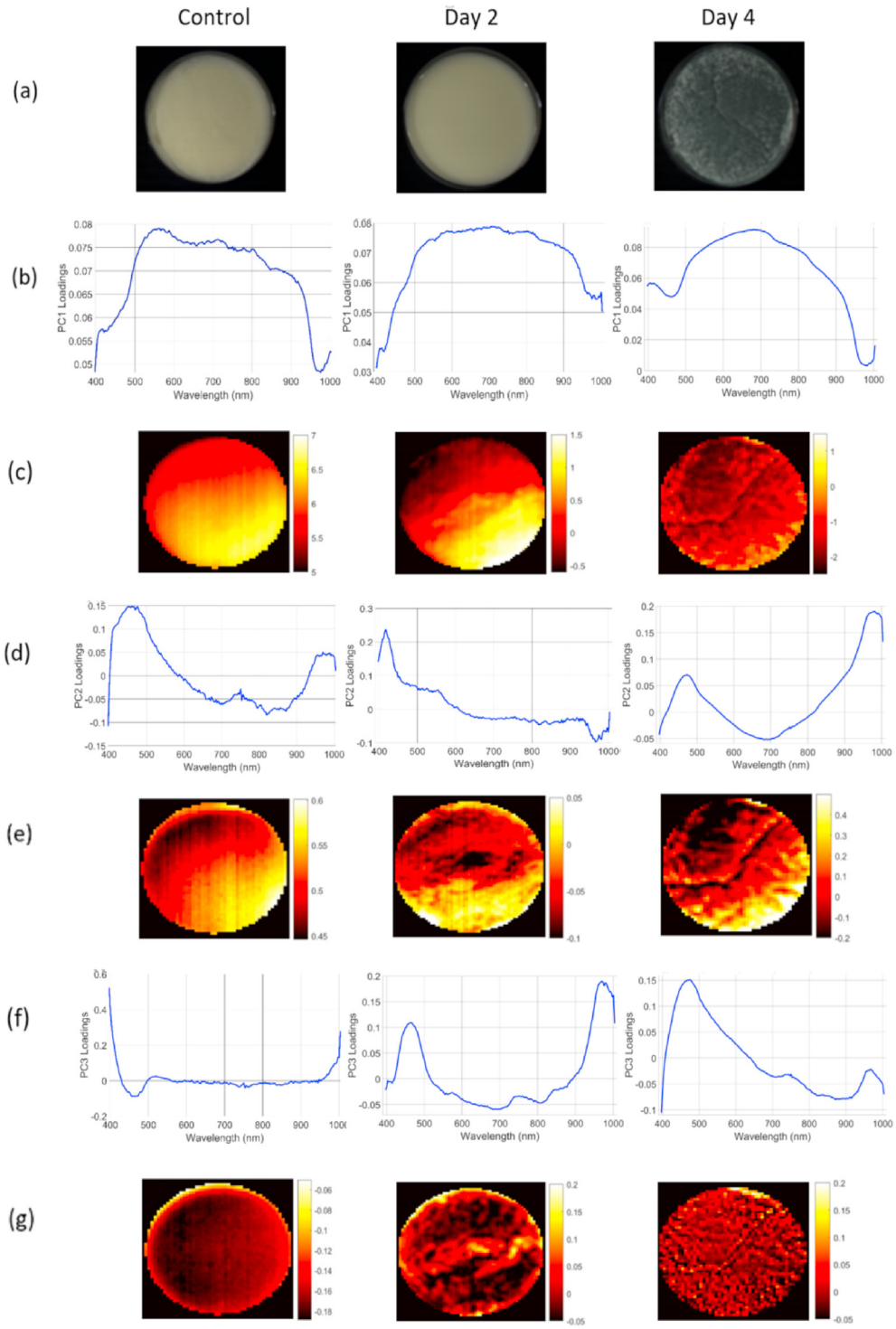

Fig. 3. Three stages (control, mould at mycelial growth and sporulation) of an agar sample showing the raw image (a), the first three PCs (b), (d) and (f) with the corresponding score image (c), (e) and (g) below each component.

corresponding hyperspectral data. The cheeselet images taken on day 1 and day 2 contain mycelial growth and sporulated areas respectively.

Similarly to the agar samples, all 10 samples that had mycelial growth or sporulated areas showed a similar spectral representation across the whole spectrum in the first three PCs. The first PC spectral representation for the control sample and the sample containing mycelia is very similar throughout the whole spectrum. However, the first PC spectral representation for the sample containing sporulated areas is different in the 400 nm–500 nm range and in the 800 nm–1000 nm range (Fig. 4b). The spectral representation of the second PC plots for the sample with mycelial growth is totally different when compared to the control sample and samples having sporulated areas (Fig. 4d). The spectral representation for the third PC loading values for the three cases of cheeselet is more of an oscillatory one.

Similar to the agar samples, the score images associated with the first PCs for the cheeselets, capture the contribution of the general illumination gradient effect of the samples, highlight shadows and bright spots (Fig. 4c). The score images associated with the second PC highlight

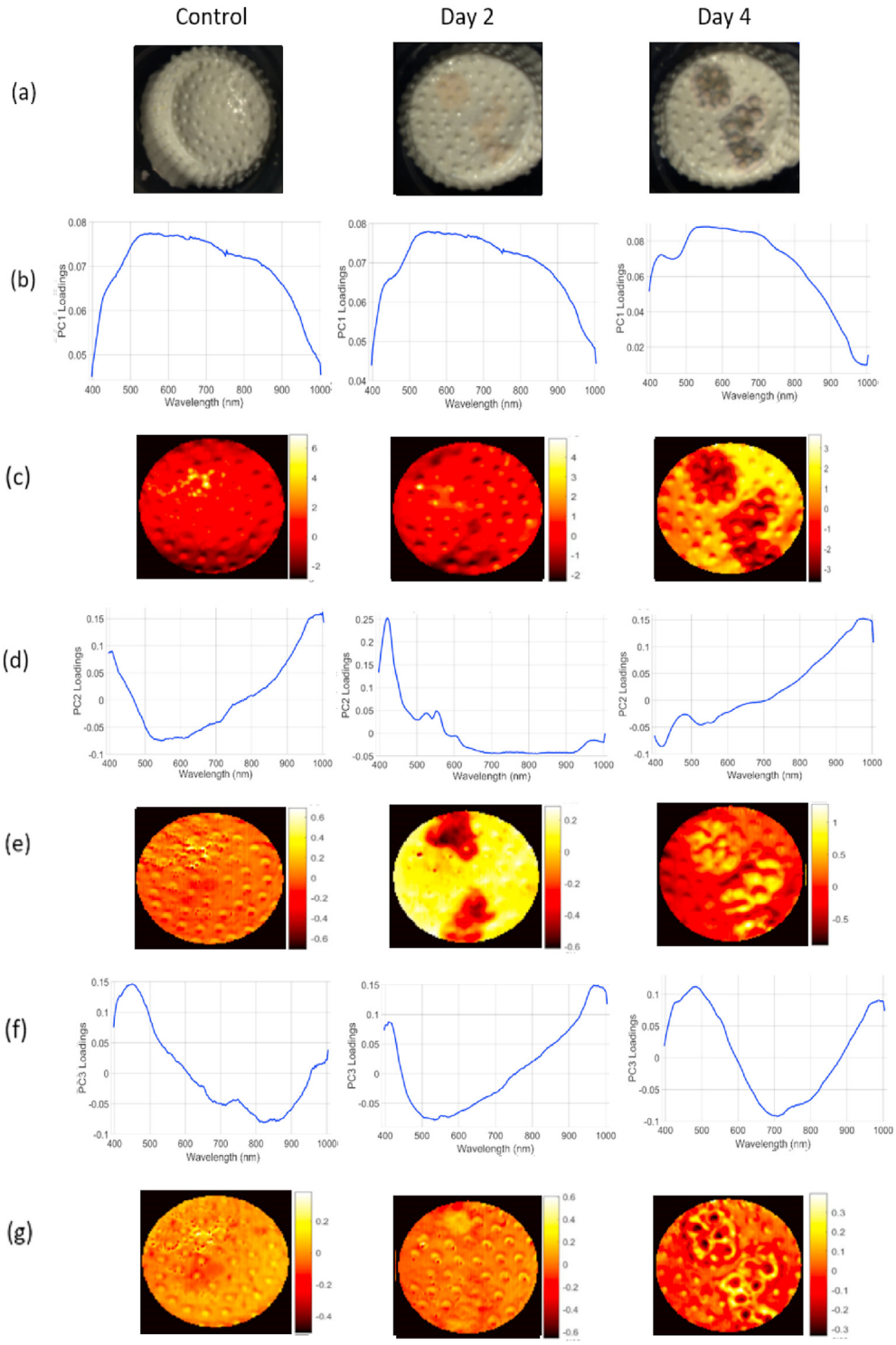

Fig. 4. Three stages (control, mould at mycelial growth and sporulation) of a cheeselet showing the raw image (a), the first three PCs (b), (d) and (f) with the corresponding score image (c), (e) and (g) below each component.

mouldy areas on the samples especially areas containing mycelia (Fig. 4e). The score images associated with third PC highlight the texture of the cheeselet surface such as the dimples as well as the border of sporulated areas (Fig. 4g). These characteristics were observed in all the score images for the cheeselets.

Fig. 5 shows the average spectral signatures and standard deviations of the ten contaminated cheeselets for the first three PCs. Two areas were selected from each of the 10 samples, one containing no contaminants and the other area containing mycelial growth. These are represented by the blue plot and the red plot in Fig. 5, respectively. The spectral signature for PC2 showed the greatest differences between the background and contaminated area when compared to the other PCs.

3.3. Results using predetermined transformation loadings

3.3.1. Testing on agar samples

Since the second component was determined to carry the most information on contaminated and uncontaminated regions of agar samples,

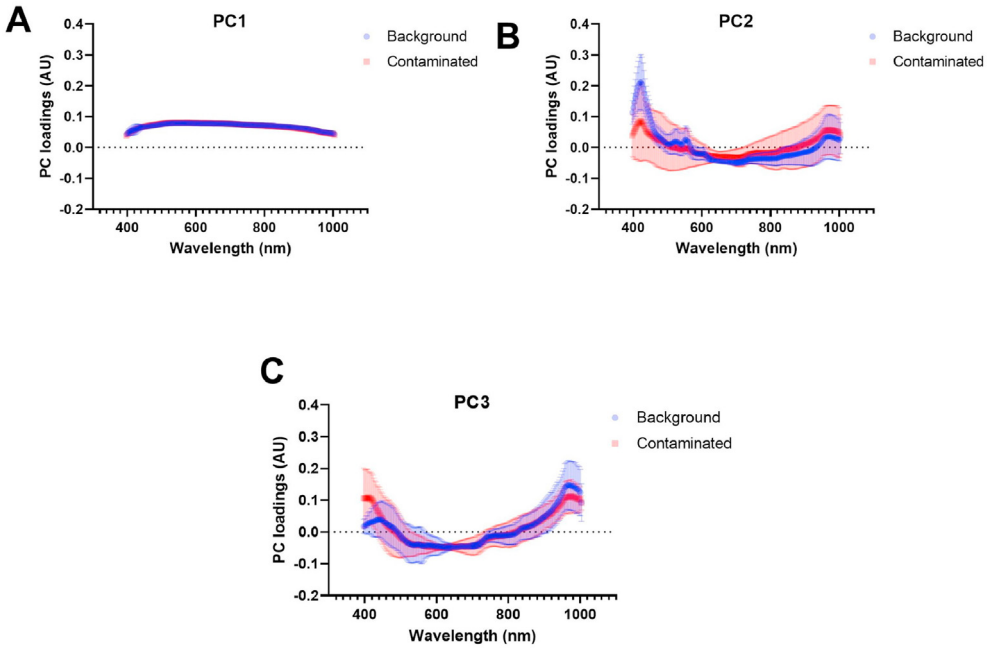

Fig. 5. Average spectral signatures with standard deviation of the 10 cheeselets for the first three PCs. The blue plot represents the background (uncontaminated) cheeselet region whilst the red plot represents a region containing mycelial growth. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

this was the PC considered for further analysis.

Fig. 6 shows the second PC for the agar samples containing mycelial growth. The blue plot shows the spectral loadings obtained from an agar sample, which were subsequently used to detect the presence of contaminants on other test samples. One may observe that all samples containing mycelia exhibit a similar behaviour. In the spectral representations of the second PC, there are some variations in the 450 nm–500 nm range and in the 900 nm–1000 nm range.

The resulted transformation coefficients from the agar sample that is highlighted in blue in Fig. 6 were multiplied to the raw data of other contaminated samples (test samples). Projected score images were computed to detect the presence or otherwise of mycelia. Fig. 7 shows an example of two agar samples containing mycelial growth. Fig. 7a contains the images that result when applying the predetermined transformation coefficients to two test samples. Fig. 7b represents the score images for the second PC obtained after applying the PCA directly on the raw data.

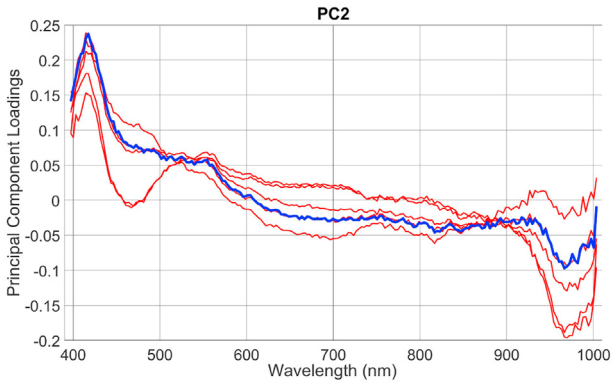

Fig. 6. Second principal component vs wavelength for agar samples containing mycelial growth. The blue plot represents an example of one of the fixed loadings that has spectral representation as the other samples. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

The use of transformation coefficients obtained from the training data were found to be effective for determining the location of mycelial growth in different test samples. A similar outcome was observed for all agar samples containing mycelia. Furthermore, when the transformation coefficients were applied to control samples, no false detections of mouldy regions were obtained.

3.3.2. *Testing on cheeselets*

Similarly to agar samples, the second PC was determined to carry the most information on contaminated and uncontaminated regions of cheeselets thus, this was the PC used for further analysis. Fig. 8 shows the second PC for the cheeselets containing mycelial growth. The blue plot

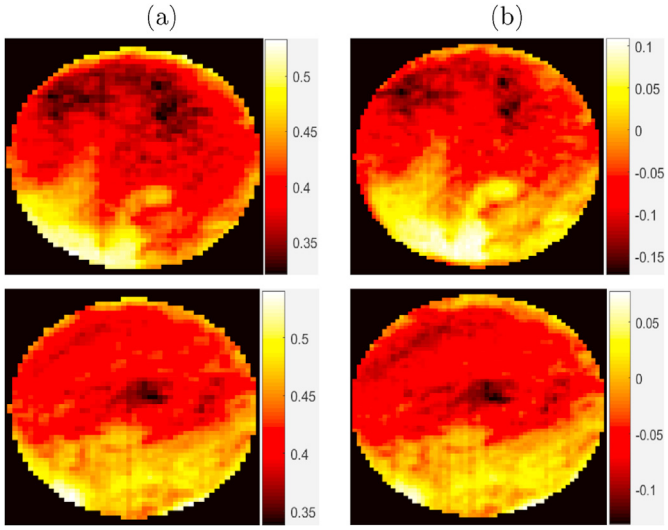

Fig. 7. The two images on the left show examples of score images for two different agar samples resulting from the application of predetermined transformation coefficients to two test samples. The two images on the right show score images of the same two samples as a result of the direct application of PCA on the raw data.

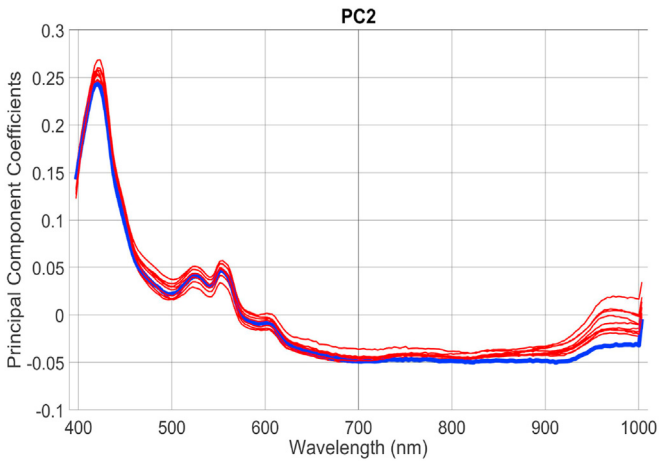

Fig. 8. Second principal component vs wavelength for cheeselets containing mycelia. The blue plot represents an example of one of the fixed loadings that has similar spectral representation as the other samples. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

shows the spectral coefficients obtained from a cheeselet which were subsequently used to detect contaminants on other samples. One may observe that all samples containing mycelia exhibit a similar behaviour. The spectral representations for different cheeselets containing mycelial growth (Fig. 8) agree better than those for agar samples (Fig. 6).

Similarly to the agar samples, the transformation coefficients obtained from the cheeselet highlighted in blue in Fig. 8 were multiplied to raw data of other contaminated samples (test samples). Projected score images were computed to detect the presence or otherwise of mycelial growth. Fig. 9 shows an example of two cheeselets containing mycelia. Fig. 9a contains the images that result when applying the predetermined transformation coefficients to two test samples. Fig. 9b represents the score images for the second PC obtained after applying the PCA directly on the raw data.

The use of transformation coefficients obtained from the training data were found to be effective for highlighting mycelia in different cheeselet test samples. It was also observed that the detected mycelia had similar area (comparing Fig. 9a and b). A similar outcome was observed for all cheeselets containing mycelial growth. Moreover, when the transformation coefficients were applied to control cheeselets, no false areas of contamination were detected.

4. Discussion

4.1. PCA on agar samples

As shown in Fig. 3b, when applied to agar samples, the first PC loadings obtained from PCA exhibit positive coefficients for all the samples, in contrast with the coefficients for the second PC loading (Fig. 3d) and the third PC loadings (Fig. 3f). Comparing the spatial projection (score images) of the first PC for the uncontaminated and contaminated agar samples, one can note that in the Day 4 sample, the sporulated areas are clearly highlighted. The first PC loadings for sporulated samples (Fig. 3b) have higher values in the 400 nm–500 nm range when compared to the control sample.

For all agar samples containing mycelia, the second PC loading values (Day 2, Fig. 3d) decrease in the 900 nm–1000 nm range and increase in the 400–450 nm range when compared to those of the control agar samples (Fig. 3d Control) and agar samples containing sporulated areas (Fig. 3d Day 4). This indicates that the second PC loading values vary when mycelial growth is present for the mentioned wavelength ranges. Moreover, the pixel values for the second PC (Fig. 3e Day 2) always have an inverted polarity for pixels containing mycelia when compared to the remaining non-contaminated agar area.

The projected images associated with the third PC (Fig. 3g) appear to highlight the texture of the agar surface. While Fig. 3 shows the results obtained from one agar sample, similar results were consistently observed across all the considered agar samples.

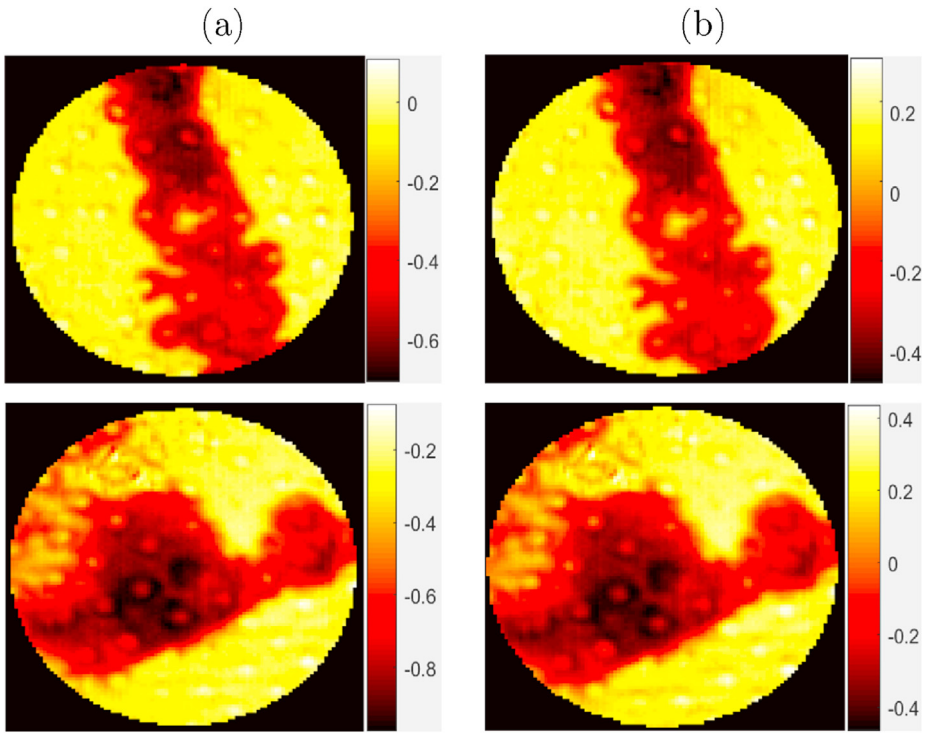

Fig. 9. The two images on the left show examples of score images for two different cheeselets resulting from the application of predetermined transformation coefficients to two test samples. The two images on the right show score images of the same two samples as a result of the direct application of PCA on the raw data.

4.2. PCA on cheeselets

Similarly to the agar samples, while the first PC loading values (Fig. 4b) are all positive for all cheeselets, this is not the case for the second (Fig. 4d) and third PC loadings (Fig. 4f). The first PC appears to be associated with the general illumination effect on the cheeselet surface. Comparing the first PC loadings for the control sample (Fig. 4b Control) to that of the sporulated samples, higher peaks are observed in the 400 nm–500 nm range for the sporulated samples (Fig. 4b Day 4). Sporulated areas on cheeselets are clearly visible on the first PC (Fig. 4c Day 4), whilst the third PC highlights the different intensities within the sporulated area as well as the border of the *P. chrysogenum* growth (Fig. 4g Day 4).

The spatial projection of the second PCs highlights areas of mycelial growth on cheeselets (Fig. 4e Day 2). The pixel values for the PC always have an inverted polarity for pixels containing mycelial growth when compared to the remaining non-contaminated cheeselet area. The difference of the second PC plot for the cheeselets containing mycelia when compared to that obtained for the control sample and the samples containing sporulated areas appears to indicate that the second PC effectively highlights the presence of mycelia (Fig. 4d). When considering the three PCs and their spatial projections, only the second PC highlights mycelial growth. In the studies done by Jiang et al. (2016) and Feng et al. (Feng and Sun, 2013), it was also stated that the second PC contained useful mouldy information on peanuts and chestnuts respectively. The characteristics discussed in this section were observed across all cheeselets.

Furthermore, the spectral signatures in Fig. 5 show that the largest differences between the PC loadings of the background area and the contaminated area occur for PC2. As was observed in Figs. 4d and 5B also shows that there is a greater change in activity for the shorter wavelengths (400 nm–500 nm range), also indicating that PC2 contains most information on mycelial growth on cheeselets and that mycelial growth can be identified by hyperspectral imaging.

4.3. Predetermined transformation loadings

The results presented in Sections 3.3.1 and 3.3.2 show that the transformation coefficients determined from training samples can be effectively used to determine the presence or otherwise of contaminants on unseen samples. Similar results were obtained when different training and test sets were used. This indicates that the PCA results associated with the contaminants from the different samples are consistent.

One of the main advantages of this approach is that in practice it would only be necessary to train the PCA on one or few samples and subsequently use the obtained loadings to test for specific contaminants on new samples being tested. This approach is appealing for scenarios where it may be of interest to train on contaminated samples with different known contaminants, obtaining predetermined loadings for every contaminant and applying these to new samples being tested to check for the presence, if any, of the contaminants.

The proposed PCA method in this study worked effectively in spatially localising and visualizing mycelial growth at the pixel level on cheeselets. While PCA has been used in a number of studies involving the use of hyperspectral imaging for food inspection, the method has primarily been employed as a preprocessing step to reduce the dimensionality of the considered data. In these cases PCA allowed for a considerable reduction of the data to be considered for further processing while retaining the most relevant information (Wang et al., 2015; MA et al., 2014).

In some cases PCA has also been applied in a manner similar to the one adopted in this work, either to separate foreground pixels of a sample of interest from background pixels, and/or to determine the contamination or otherwise of food products such as peanuts and maize kernels (Jiang et al., 2016; Foca et al., 2016).

The effectiveness of a pixel-wise application of PCA for the detection of fungal contamination in cheeselets was shown in this study. Moreover, it was shown that the evolution of fungal contamination can also be determined through the resulting PCs. Since hyperspectral imaging can provide a fine representation of the spectral signature of the product or contaminant under investigation, it is plausible to assume that the proposed approach would also be effective on different types of cheeses and contaminants. It is also worth noting that while the spectral range of the hyperspectral imaging equipment used in this work overlaps considerably with the visible spectrum captured by standard RGB cameras, any fine differences in the spectral signatures of the considered food products or contaminants are expected to only be discernible thanks to the fine spectral resolution that a hyperspectral imaging system offers. It is in fact expected that the use of the PCA approach in conjunction with hyperspectral imaging, as considered in this work, would be suitable to discriminate between contaminants even in cases of minor differences in their spectral signatures. However, further testing would be required to confirm this in a practical setting.

5. Conclusion

In this study, an analysis was proposed to detect the presence of mycelial growth on milk agar and cheeselets samples. The score images obtained from the proposed PCA procedure indicated that the second PC of contaminated samples was consistently representative of the presence of contaminants on the inspected sample.

It was also proposed the testing of cheeselets for contamination by determining PCA loadings that were obtained from training samples that could subsequently be used to determine the presence of contaminants in test samples. This approach was found to be effective in localising the presence of mycelia on both the agar and cheeselets.

The proposed analysis may provide a rapid, non-contact food inspection approach through the use of hyperspectral imaging in food production chains. While this study focused on the detection of contaminants in cheeselets, the same concepts could be adapted and applied to other solid and semi-solid food products.

Declaration of competing interest

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

Acknowledgements

The ‘Food Inspection using Hyperspectral Imaging’ (FIHI) project is financed by the Malta Council for Science and Technology, for and on behalf of the Foundation for Science and Technology, through the FUSION R&I technology development programme (MCST R&I-2015-048-T). The project was also support by COST Action CA15118, "Mathematical and Computer Science Methods for Food Science and Industry

References

Andersen, B., Frisvad, J.C., Søndergaard, I., Rasmussen, I.S., Larsen, L.S., 2011. Associations between fungal species and water-damaged building materials. Appl. Environ. Microbiol. 77, 4180–4188.

Barbin, D., Elmasry, G., Sun, D.W., Allen, P., Morsy, N., 2013. Non-destructive assessment of microbial contamination in porcine meat using NIR hyperspectral imaging. Innovat. Food Sci. Emerg. Technol. 17, 180–191. https://doi.org/10.1016/j.ifset.2012.11.001.

Barreto, A., Cruz-Tirado, J., Siche, R., Quevedo, R., 2018. Determination of starch content in adulterated fresh cheese using hyperspectral imaging. Food Biosci. 21, 14–19.

Burke, N., Zacharski, K.A., Southern, M., Hogan, P., Ryan, M.P., Adley, C.C., 2018. The dairy industry: process, monitoring, standards, and quality. In: Descriptive Food Science (IntechOpen).

Calvini, R., Michelini, S., Pizzamiglio, V., Foca, G., Ulrici, A., 2020. Exploring the potential of nir hyperspectral imaging for automated quantification of rind amount in grated parmigiano reggiano cheese. Food Contr. 112, 107111.

Del Fiore, A., Reverberi, M., Ricelli, A., Pinzari, F., Serranti, S., Fabbri, A., Bonifazi, G., Fanelli, C., 2010. Early detection of toxigenic fungi on maize by hyperspectral imaging analysis. Int. J. Food Microbiol. 144, 64–71.

Feng, L., Zhu, S., Lin, F., Su, Z., Yuan, K., Zhao, Y., He, Y., Zhang, C., 2018. Detection of oil chestnuts infected by blue mold using near-infrared hyperspectral imaging combined with artificial neural networks. Sensors 18, 1944.

Feng, Y., Sun, D.W., 2012. Application of hyperspectral imaging in food safety inspection and control: a review. Crit. Rev. Food Sci. Nutr. 52, 1039–1058. https://doi.org/10.1080/10408398.2011.651542.

Feng, Y., Sun, D.W., 2013. Determination of total viable count (TVC) in chicken breast fillets by near-infrared hyperspectral imaging and spectroscopic transforms. Talanta 105C, 244–249. https://doi.org/10.1016/j.talanta.2012.11.042.

Foca, G., Ferrari, C., Ulrici, A., Sciutto, G., Prati, S., Morandi, S., Brasca, M., Lavermicocca, P., Lanteri, S., Oliveri, P., 2016. The potential of spectral and hyperspectral-imaging techniques for bacterial detection in food: a case study on lactic acid bacteria. Talanta 153, 111–119. https://doi.org/10.1016/j.talanta.2016.02.059.

Gowen, A., O'Donnell, C., Burger, J., O-Callaghan, D., 2011. Analytical Methods | Hyperspectral Imaging for Dairy Products, pp. 125–132. https://doi.org/10.1016/B978-0-12-374407-4.00513-6.

Gowen, A.A., Burger, J., O'Callaghan, D., O'Donnell, C., 2009. Potential applications of hyperspectral imaging for quality control in dairy foods. In: 1st International Workshop on Computer Image Analysis in Agriculture. Potsdam, Germany.

Griffin, S., Falzon, O., Camilleri, K., Valdramidis, V.P., 2020. Bacterial and fungal contaminants in caprine and ovine cheese: a meta-analysis assessment. Food Res. Int. 137, 109445.

Huang, L., Zhao, J., Chen, Q., Zhang, Y., 2013. Rapid detection of total viable count (TVC) in pork meat by hyperspectral imaging. Food Res. Int. 54, 821–828.

Jiang, J., Qiao, X., He, R., 2016. Use of near-infrared hyperspectral images to identify moldy peanuts. J. Food Eng. 169, 284–290.

Kamruzzaman, M., Makino, Y., Oshita, S., 2015. Non-invasive analytical technology for the detection of contamination, adulteration, and authenticity of meat, poultry, and fish: a review. Anal. Chim. Acta 853, 19–29.

Lei, T., Lin, X.H., Sun, D.W., 2019. Rapid classification of commercial cheddar cheeses from different brands using plsda, lda and spa–lda models built by hyperspectral data. J. Food Meas. Char. 13, 3119–3129.

Lei, T., Sun, D.W., 2019. Developments of nondestructive techniques for evaluating quality attributes of cheeses: a review. Trends Food Sci. Technol. 88, 527–542.

Ma, T., A, M., L, R., Jayas, D., 2014. Near infrared (nir) hyperspectral imaging to classify fungal infected date fruits. J. Stored Prod. Res. 59, 306–313.

McLean, S., Dunn, L., Palombo, E., 2010. Applications of polymerase chain reaction in the dairy industry. Aust. J. Dairy Technol. 65, 81–85.

Munir, M., Wilson, D.I., Yu, W., Young, B., 2018. An evaluation of hyperspectral imaging for characterising milk powders. J. Food Eng. 221, 1–10.

Parseelan, A., Muthu, S., Kannan, P., Ayyasamy, E., Narayanan, R., 2019. Aerobic plate count of milk and dairy products marketed in different zones of Chennai. Int. J. Livestock Res. https://doi.org/10.5455/ijlr.20180711104721.

Samson, R., Houbraken, J., Thrane, U., Frisvad, J., Andersen, B., 2010. Food and Indoor Fungi. CBS-KNAW Fungal Biodiversity Centre, Utrecht, The Netherlands, p. 390.

Shan, J., Zhang, Y., Liang, J., Wang, X., 2020. Characterization of the processing conditions upon textural profile analysis (tpa) parameters of processed cheese using near-infrared hyperspectral imaging. Anal. Lett. 53, 1190–1203.

Sun, D.W., 2010. Hyperspectral Imaging for Food Quality Analysis and Control. Elsevier.

Tao, F., Peng, Y., Gomes, C.L., Chao, K., Qin, J., 2015. A comparative study for improving prediction of total viable count in beef based on hyperspectral scattering characteristics. J. Food Eng. 162, 38–47.

Wang, W., Heitschmidt, G., Windham, W., Feldner, P., Ni, X., Chu, X., 2014. Feasibility of detecting aflatoxin b1 on inoculated maize kernels surface using Vis/NIR hyperspectral imaging. J. Food Sci. 80 https://doi.org/10.1111/1750-3841.12728.

Wang, W., Ni, X., Lawrence, K.C., Yoon, S.C., Heitschmidt, G.W., Feldner, P., 2015. Feasibility of detecting aflatoxin b1 in single maize kernels using hyperspectral imaging. J. Food Eng. 166, 182–192.