



**Polak, Adam and Coutts, Fraser Kenneth and Murray, Paul and Marshall, Stephen (2019) The use of hyperspectral imaging for cake moisture and hardness prediction. IET Image Processing. pp. 1-10. ISSN 1751-9659 , <http://dx.doi.org/10.1049/iet-ipr.2018.5106>**

This version is available at <https://strathprints.strath.ac.uk/67257/>

**Strathprints** is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (<https://strathprints.strath.ac.uk/>) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to the Strathprints administrator: [strathprints@strath.ac.uk](mailto:strathprints@strath.ac.uk)

The Strathprints institutional repository (<https://strathprints.strath.ac.uk>) is a digital archive of University of Strathclyde research outputs. It has been developed to disseminate open access research outputs, expose data about those outputs, and enable the management and persistent access to Strathclyde's intellectual output.

# The use of Hyperspectral Imaging for Cake Moisture and Hardness Prediction

Adam Polak<sup>1,2\*</sup>, Fraser K. Coutts<sup>1</sup>, Paul Murray<sup>1</sup>, Stephen Marshall<sup>1</sup>

<sup>1</sup> Centre for Signal and Image Processing, University of Strathclyde, 204 George Street, Glasgow, UK

<sup>2</sup> Fraunhofer Centre for Applied Photonics, Fraunhofer UK Research Ltd., 99 George Street, Glasgow, UK

\*[adam.polak@strath.ac.uk](mailto:adam.polak@strath.ac.uk)

**Abstract:** Industrial baking of sponge cakes requires various quality indicators to be measured during production such as moisture content and sponge hardness. Existing techniques for measuring these properties require randomly selected sponges to be removed from the production line before samples are manually cut out of each sponge in a destructive way for testing. These samples are subsequently processed manually using dedicated analysers to measure moisture and texture properties in a lengthy process, which can take a skilled operator around 20 minutes to complete per sponge. In this paper, we present a new, single sensor hyperspectral imaging approach, which has the potential to measure both sponge moisture content and hardness simultaneously. In the last decade, hyperspectral imaging systems have reduced in cost and size and, as a result, they are becoming widely used in a number of industries and research areas. Recently, there has been an increased use of this technology in the food industry and in food science applications and research. The application of this technology in the cake production environment, empowered by sophisticated signal and image processing techniques and prediction algorithms as presented in this paper has the potential to provide on-line, real-time, stand-off cake quality monitoring.

## 1. Introduction

Food production is constantly confronted with increasingly rigorous requirements on the product quality, ingredients, and shelf lifetime. In this age of mass food production, which is facilitated by a network of retailers enabling global distribution, product shelf life has become one of the critical quality factors that manufacturers must maximise and accurately predict to position them strongly in a highly competitive market. Various metrics are used to assess and track the quality of baked goods during production, but these are often based on various statistical sampling schedules and the measurement procedures themselves are often time-consuming, expensive, and destructive. The proposal of an on-line, stand-off test, monitoring 100% of produced goods is a solution that has the potential to solve many quality control issues.

Hyperspectral Imaging (HSI) is a spectroscopic imaging technique which allows spatial and spectral images of a scene to be acquired. Unlike conventional cameras which sample red, green and blue regions of the spectrum, hyperspectral imaging systems capture hundreds or thousands of images, each at a different wavelength of the electromagnetic spectrum. Historically, due to the cost of HSI systems, most applications were in remote sensing [1] and wide area surveillance [2]. In recent years, the reduction in cost and size of HSI systems has made them increasingly more accessible for everyday applications in industry and research; furthermore, HSI is one of a number of modern technologies that has found application in on-line product monitoring. The inherent spectral data acquired by these systems, combined with appropriate chemometric analysis can provide a powerful tool for quality control. On-line applications of HSI range from the mining industry [3]–[5], through to the textile, pharmaceuticals [6]–[10], and many other key industry sectors. The application of HSI also has enormous potential in the food industry [3], [11]–[14]; here, the most common on-line applications of HSI are for the

inspection and segregation of agricultural crops and produce [13], [15], [16]. A number of additional cases of HSI inspection of food production processes – including the inspection of baked products – have also been presented in the literature [17]–[21].

HSI inspection can be used for segregation of different types of products, such as different types of grains, including maize, wheat, barley, oat, soybean, and rice seed [16], identification of ripeness [14], [22], and detection of flaws – such as bruising in fruit [23], [24]. It can also be used for the indication of nutritive properties, such as fat and protein content [25]–[27]. Another parameter that can be directly measured by HSI is moisture content [28]. The level of moisture in a product under inspection is an important parameter not only for the food industry, but also many others such as paper manufacturing or wood drying processes [29]–[31]. Monitoring of the moisture for food products has been reported in [14], [16], [32] and work specifically targeting the analysis of baked products has also been presented in [21], [33], [34], and [35]. In [34], Whitworth et al. present a number of applications of HSI in the analysis of food including techniques for the detection of added gluten in wheat and for studying fat distribution in French fries. The authors also study the evolution of moisture in baguettes when stored at different levels of humidity. In fact, in recent years, HSI has been used in the baking industry to analyse a number of products, including bread, biscuits, cakes, and other goods for the purposes of product innovation and reformulation and for making healthier alternatives [35]. In [21], the authors use spectral imaging to characterise the quality of butter cookies based on a so-called browning score. The browning score is computed from the spectral images and varies with the oven temperature used when baking the cookies and with the bake time. The results in [21] are shown to correlate well with scores allocated by a tasting panel, who scored the cookies based on their quality.

In this paper we present a novel application of HSI for analysing baked products, with a focus on addressing the

challenge of simultaneously measuring both the moisture content and hardness of baked sponges in a stand-off and efficient way. While various authors have presented the use of HSI to analyse the properties of baked goods, to the best of the authors' knowledge, a single system capable of simultaneously measuring both moisture content and sponge hardness has not been presented. To support and justify the use of HSI in this context, we investigate the robustness of using reflectance spectra to quantitatively predict cake hardness and moisture through comparison with ground-truth data. We also explore the impact that the spectral data range and spectral resolution has on the prediction model performance, since it is possible that a full hyperspectral image may not be required for this application. This allows us to explore how well a multispectral or even single band system could be designed to create a new, fast, effective, and affordable moisture and hardness monitor for the baking industry and beyond.

## 2. Data Acquisition

Hyperspectral images of a large number of sponges of different ages were acquired alongside ground-truth moisture and hardness measurements for each sponge, which were obtained using traditional destructive techniques. To allow efficient data capture, a total of 96 circular sponge cakes – of which 48 were composed of white, plain sponge and 48 were composed of brown, chocolate sponge – were collected and stored immediately after baking over a period of six weeks. At the end of the six week period, the age of the collected sponges – which each had a diameter of 9" – ranged from a few hours to 45 days old; this allowed hyperspectral image, moisture, and hardness data for a large age range of sponges to be collected in just a few days.

Prior to imaging, the top of each sponge was removed such that the height from the base to the surface of each sponge measured exactly 3 cm in height. This was achieved using a specially designed fixture to ensure accuracy and repeatability between sponges. An image of a sponge already prepared for imaging is shown in Figure 1. Each sponge cake was then imaged using the GmbH, Red Eye 1.7 near-infrared (NIR) system which acquires HSI data sampled at 256 discrete wavelengths in the spectral range 900 nm – 1700 nm. Of particular note within this range is a strong water (moisture) absorption band at approximately 1450 nm [28]. As a passive system, the Red Eye 1.7 requires external illumination. To ensure a sufficient level of infrared radiation, a set of two 12 V DC off-the-shelf halogen lamps were used during imaging. The halogen lamps, as an incandescent light source, provide excellent illumination in the near-IR range, as required by this system [36]. All hyperspectral images had spatial dimensions of 320-by-400 and a bit-depth of 14 bits per pixel. Prior to the imaging of each sponge, a flat tile of Spectralon® (a material of high Lambertian reflection over its reflective spectral range of 250 nm – 2500 nm [37]) was placed at the height of the sponge surface and used to calibrate the HSI camera's white levels, and the lens cap was applied to calibrate the black levels.

Following acquisition of the HSI data for a sponge, a circular stencil with a diameter of 3 cm was used to cut a sample from the centre of the sponge for hardness and moisture analysis. To measure hardness, each sample was analysed using a Brookfield CT3 Texture Analyser. This



**Fig. 1.** Use of sponge cutting fixture and 3 cm circular stencil to prepare a chocolate sponge.

process irreversibly changed the texture parameters of the sponge, but did not affect the moisture, therefore the same sample could be used for moisture analysis. For this, 4 g of cake crumbs were placed in a Mettler Toledo HB43-S Halogen Moisture Analyser and the moisture content of the sponge was recorded. After data acquisition, each sample was disposed of and the testing equipment was cleaned to remove any residual sponge material and the procedure was repeated until all the collected cakes had been processed.

The sponge moisture metric was equal to the percentage of the 4 g sponge sample mass loss upon drying under a halogen lamp. The sponge texture metrics, as provided by the texture analyser, were Hardness 1, Hardness 2, Cohesiveness, Springiness, and Adhesiveness. While measurements were made for all of the texture metrics, only measurements of Hardness 1 and Hardness 2 were used in the development of the prediction model since these were the only ones shown to correlate well with moisture as will be explained in Section 3. Hardness 1 and Hardness 2 are equal to the peak loads, measured in grams, of the first and second cake sample compression cycle, respectively [38].

After measuring all 96 sponges using the methods described above, the resulting data set collected for each cake contained a calibrated set of predictors (the calibrated reflectance data from the Red Eye 1.7 NIR HSI system) and responses (the physical properties of interest – moisture and texture) which could be used to build an appropriate prediction model to allow moisture and hardness to be predicted directly from hyperspectral images of baked sponge.

It is worth noting that there are water absorption lines in the Vis-NIR range (400 nm – 1000 nm), where significantly less costly silicon based systems could be used, and there are successful examples of this in the literature [39]. However, these lines are much weaker; even the strongest at approximately 970 nm falls at the sensitivity edge of these detectors, where signal-to-noise ratio is relatively low. While initial results using a Vis-NIR system, targeting the water line at 970 nm were promising [11], during our experiments, we found the 1450 nm band most useful for this application.

		Cell Contents: Pearson correlation P-Value					
<b>Results for Type = Chocolate</b>							
	Hardness 1 [g]	Hardness 2 [g]	Cohesiveness	Springiness [mm]	Adhesiveness [mJ]	Moisture @ centre	
Hardness 2 [g]	0.994 0.000						
Cohesiveness	-0.565 0.000	-0.503 0.000					
Springiness [mm]	-0.413 0.004	-0.394 0.006	0.350 0.015				
Adhesiveness [mJ]	0.033 0.825	-0.008 0.958	-0.345 0.016	-0.143 0.331			
Moisture @ centre	-0.774 0.000	-0.724 0.000	0.672 0.000	0.428 0.002	-0.189 0.199		
Reflectance @ 1450nm	0.646 0.000	0.584 0.000	-0.539 0.000	-0.389 0.006	0.138 0.349	-0.792 0.000	
<b>Results for Type = White</b>							
	Hardness 1 [g]	Hardness 2 [g]	Cohesiveness	Springiness [mm]	Adhesiveness [mJ]	Moisture @ centre	
Hardness 2 [g]	0.999 0.000						
Cohesiveness	-0.327 0.023	-0.293 0.044					
Springiness [mm]	-0.312 0.031	-0.299 0.039	0.360 0.012				
Adhesiveness [mJ]	0.411 0.004	0.399 0.005	-0.394 0.006	-0.240 0.100			
Moisture @ centre	-0.902 0.000	-0.901 0.000	0.289 0.046	0.274 0.060	-0.427 0.002		
Reflectance @ 1450nm	0.830 0.000	0.826 0.000	-0.316 0.028	-0.305 0.035	0.409 0.004	-0.960 0.000	

**Fig. 2.** Results of the correlation test for chocolate and white sponge cakes separately.

### 3. Development of the Prediction Model

Before building a prediction model, the Pearson Correlation Coefficient (a measure of the linear correlation between two variables) [40], [41] was used to provide an indication of the relationship between the data sets. This established the range of sponge parameters that could be predicted using the reflectance measurements, and confirmed our understanding of the relationship between reflectance and moisture as well as hardness measurements. Once this relationship was understood, a suitable model based on Partial Least Squares Regression (PLS-R) was designed.

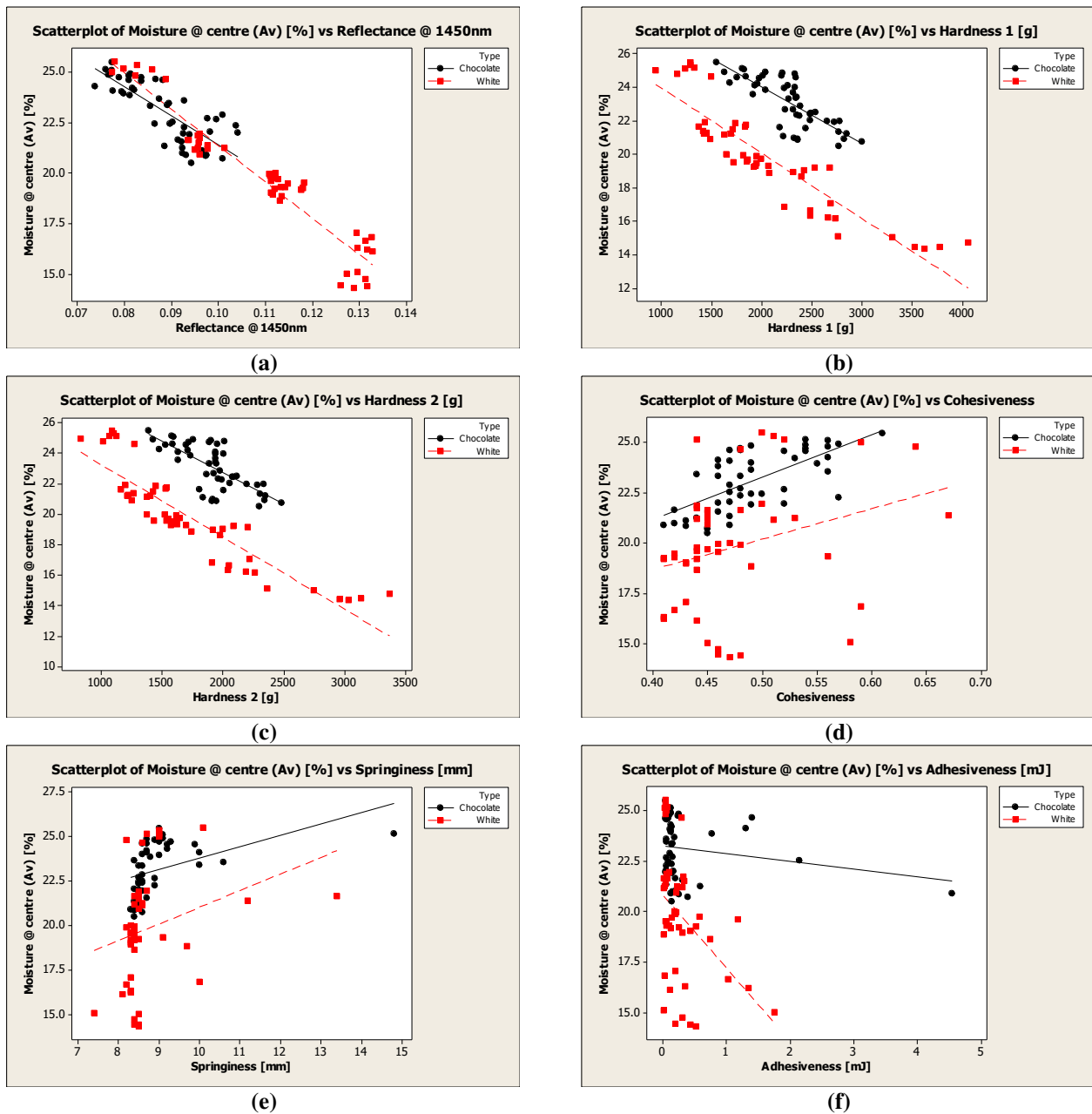
#### 3.1. Correlation Tests

It is well documented that water (moisture) has distinct absorption features in the near infrared (NIR) spectral range [28]; however, no direct links between reflectance and tactile properties are known to the authors. That said, correlation between sponge moisture and hardness does exist; as a sponge loses moisture, it becomes more firm. The Pearson Correlation Coefficient was used to check the strength of the relationship between the moisture level and the spectral data (from one chosen wavelength at 1450 nm) for both tested cake types. It was also used to establish whether or not there existed any relationship between the moisture and any of the other properties measured.

Figure 2 demonstrates the result of the aforementioned correlation test. The correlation was deemed to be: strong if the Pearson Correlation Coefficient (PCC) > 0.9 (marked with a green box); average if  $0.9 > PCC > 0.6$  (marked with the yellow box); and weak if  $PCC < 0.6$  (marked by the red box). The p-Value – noted underneath each PCC in Figure 2 – indicates the significance of the relationship computed. If

the value is small (typically less than 0.05), there is strong statistical evidence against the null hypothesis. In the case of regression analysis, this hypothesis implies that there is no relationship between independent variable  $X$  and dependent variable  $Y$ . Thus, if the p-Value is low it means that the null hypothesis is rejected, which in turn implies that there is indeed a relationship between the dependent and independent variables. In our analysis, all the characteristics with a high PCC average have a p-Value equal to zero, indicating a significant relationship for each. The correlation between reflectance, moisture, and all other texture properties measured can also be illustrated graphically, as shown in Figure 3. Here, the data is shown in a series of scatter plots with computed lines of best fit overlaid.

For chocolate sponges, there is, on average, a negative correlation between the reflectance data and moisture. The same is true for the relationship between moisture and Hardness 1 and Hardness 2. There is also a trend between cohesiveness and moisture (which is positive on average) and all of these relationships are statistically significant. For white sponges, there is strong negative correlation between the reflectance data and moisture, and between moisture and Hardness 1 and Hardness 2. While all these relationships are statistically significant, the correlation between moisture and cohesiveness is not as strong as for the chocolate sponges. The springiness and adhesiveness properties have weak correlation with moisture and were omitted in further analysis. Cohesiveness was also omitted from further analysis since correlation only exists for chocolate sponges. Overall, the strong correlation between the reflectance data and the moisture measurements provides confidence that moisture can be reliably predicted using the HSI data. Furthermore, our experiments indicate that Hardness 1 and Hardness 2 can be predicted from moisture, therefore it should be possible to use the HSI system's moisture predictions to estimate Hardness 1



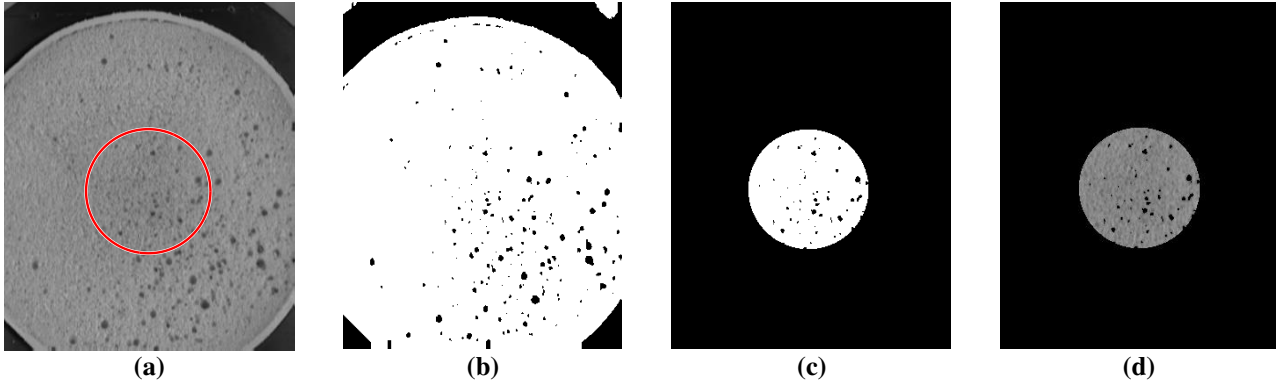
**Fig. 3.** Scatterplots with regression for analysed data relationships between (a) reflectance and moisture, (b) Hardness 1 and moisture, (c) Hardness 2 and moisture, (d) cohesiveness and moisture, (e) springiness and moisture, and (f) adhesiveness and moisture.

and Hardness 2 for both white and chocolate sponge cakes in a fast, repeatable, and stand-off way.

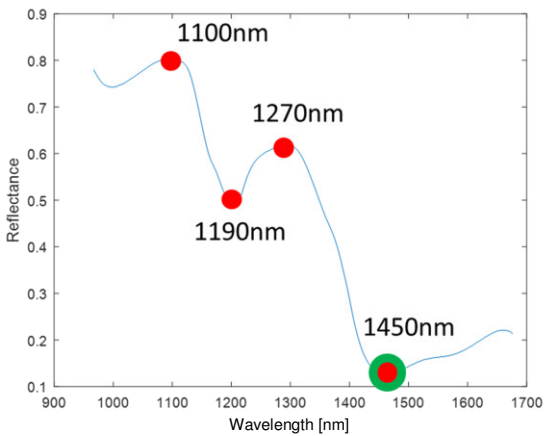
### 3.2. Partial Least Squares Regression for moisture and hardness prediction using HSI

PLS-R is a statistical method for finding a relationship between a set of predictors and responses [42], [43]. It may be used to predict a value of interest based on another correlated parameter. PLS-R exists in various implementations, but the two-block predictive version is the most popular form in science and technology [44]. This models the relationship between two data matrices,  $X$  and  $Y$ , and can analyse strongly correlated and noisy data within matrix  $X$  in order to model a number of response variables in

matrix  $Y$ . The mathematical method at the core of this modelling iteratively estimates each model parameter at the slope of a simple bivariate regression between the data of the  $y$  and  $x$  variables [42], [43]. Similar to simple regression, the two data blocks  $X$  and  $Y$  are respectively called predictors or independent variables (although, as mentioned above, often this data tends to be highly correlated and therefore the name independent is not always adequate), and responses or dependent variables, meaning the variables whose variation we wish to explain or predict. The method minimises error scores across all individuals to enhance the prediction. The test uses the principle of least squares, which selects, among many possible lines, the one that best fits the data (and minimises the sum of squared vertical distances from the observed data points to the line). With this in mind, typically



**Fig. 4.** Processing of spatial image, extracted from HSI data cube, to exclude the air bubbles from further analysis: (a) single band (1450 nm) of hyperspectral data with indication of the centre area, (b) binary image obtained by thresholding indicating location of the cavities (in black), (c) binary mask selecting the centre of the cake (white), excluding air bubbles (black), (d) selection of the cake ready for spectral data extraction.



**Fig. 5.** Illustration of the spectral data selection. Blue curve - hyperspectral data; red points - multispectral data; green circle - single band data.

a more descriptive expansion of the acronym PLS is Projection to Latent Structures [45]. As PLS-R is explicitly able to explain the variability in the responses, we have selected this method to construct prediction models to predict the characteristics of the baked sponges we have imaged. For our application, we wish to use PLS-R to predict sponge moisture content and hardness based on our acquired HSI reflectance spectra.

### 3.3. Information Extraction from HSI Data

The spectral response of each sponge was computed as the median value of all valid spectra in the central part of the cake. A full analysis of all pixels covering the sponge surface could introduce unnecessary errors, as sponge cakes can contain air bubbles and other structural defects. During imaging, structural imperfections can act as cavities that significantly affect the amount of reflected light. To avoid this potential source of error in our model, data from the surface of cavities was excluded from the analysis. Figure 4 illustrates the general steps of an automated image processing algorithm, which has been designed to detect cavities and subsequently ignore them in the analysis. The algorithm first

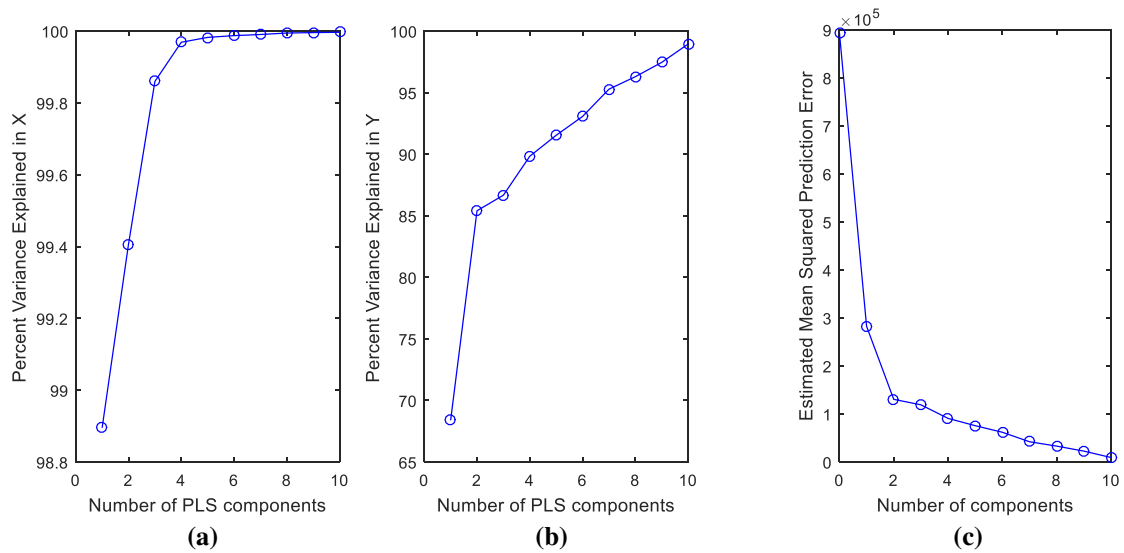
applies thresholding to identify bright regions of the image likely to correspond to the sponge. Then, the perimeter of the sponge is detected using the Hough Transform [46] for circle detection, and a circle of radius  $r$  in the centre of the detected sponge is identified and extracted for further analysis. Circular structural defects are detected using morphological techniques and removed from the analysis.

### 3.4. Spectral Data Selection

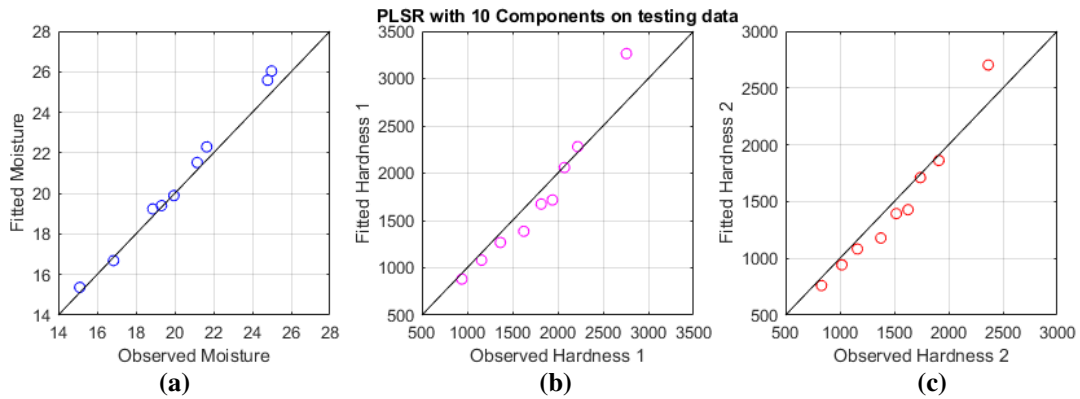
During data acquisition, reflectance data was captured in the form of hyperspectral images containing 256 spectral bands over 800 nm ranging from 900 nm to 1700 nm. However, due to decreases in SNR at the edges of the camera detector, only 225 bands were selected, leaving data from a range of wavelengths between 967 nm and 1676 nm to be used for further analysis. While hyperspectral data sets are very rich in information, they require sophisticated and relatively expensive equipment to capture the data. As a result of this, we also assess the feasibility of using a multispectral camera (which utilises only a subset of spectral bands) or a single band camera in place of a full HSI system. We do this by constructing three prediction models for each sponge type (white and chocolate), each with a different number of spectral bands contributing to the model predictors. In other words, we construct separate models to predict moisture from:

- 1) a raw, information rich hyperspectral data cube;
- 2) a simulated multispectral image obtained by subsampling the hyperspectral data cube;
- 3) a simulated single band dataset obtained by selecting one image from the hyperspectral data cube which is known to correspond to water absorption.

Figure 5 illustrates a typical reflectance spectra captured using the HSI camera (blue line), the selection of four bands simulating the multispectral dataset (red circles), and the single wavelength selected from the moisture absorption band (green circle). The simulated multispectral dataset was constructed from four discrete wavelengths, which were selected based on the fact that they represented four distinct stationary points in the spectrum. For the single image experiment, the image corresponding to 1450 nm was selected as this wavelength is at the centre of a strong moisture absorption band [28].



**Fig. 6.** Graphs illustrating (a) percent of variance explained for predictors and (b) responses, and (c) estimated Mean Square Prediction Error when varying number of PLS-R components for white cakes prediction model with HSI data as predictors.



**Fig. 7.** Graphs comparing predicted (fitted) and measured (observed) values of (a) moisture, (b) Hardness 1, and (c) Hardness 2 for a white cake PLS-R prediction model trained and tested on HSI data.

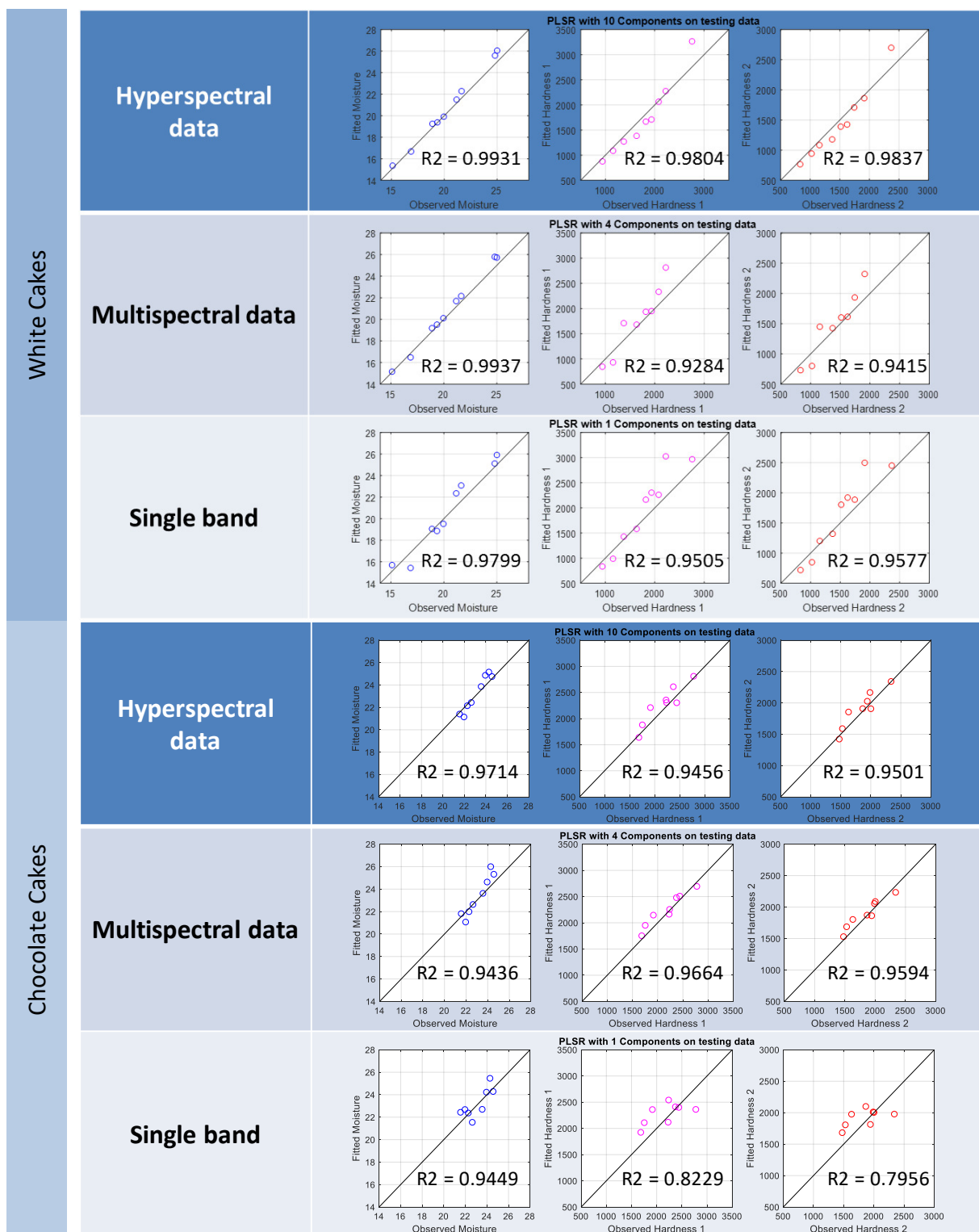
**Table 1** Table of observed and predicted values of moisture, Hardness 1, and Hardness 2, and  $R^2$  value for each prediction, for a white cake PLS-R prediction model trained and tested on HSI data.

Cake No (age)	Moisture		Hardness 1		Hardness 2	
	Predicted	Observed	Predicted	Observed	Predicted	Observed
1 (5)	25.58	24.79	1080.97	1158.50	942.53	1018.50
2 (37)	15.37	15.11	3263.29	2757.50	2702.44	2365.00
2 (12)	22.30	21.65	1267.40	1367.00	1082.45	1161.00
4 (19)	21.53	21.16	1386.00	1625.50	1178.95	1377.00
5 (27)	19.40	19.32	1716.64	1941.00	1428.30	1627.00
6 (28)	19.90	19.96	1674.13	1817.50	1394.44	1519.00
7 (41)	16.69	16.85	2280.66	2220.00	1861.83	1911.00
8 (43)	19.24	18.87	2060.46	2074.00	1711.34	1740.50
9 (0)	26.04	24.99	880.99	942.00	760.44	831.00
<b><math>R^2</math></b>	<b>0.9931</b>		<b>0.9804</b>		<b>0.9837</b>	

### 3.5. Fitting the Model

A PLS-R model can be designed to establish the relationship between two sets of data – predictors and responses. For our application, the reflectance data (from three different datasets: hyperspectral, simulated multispectral and single band) were used as the predictors and the moisture, Hardness 1, and Hardness 2 data were used as

the responses of the three models in Sec. 3.4. Each model must be trained to learn how to establish all the relationships between predictors and responses. As such, the majority (81.25%) of the acquired data was used as a training set to build each of the models. This training set contained spectral information from 39 cakes of each sponge type; these cakes were of various ages, but none of those selected were baked on the same day that hyperspectral data was captured. The



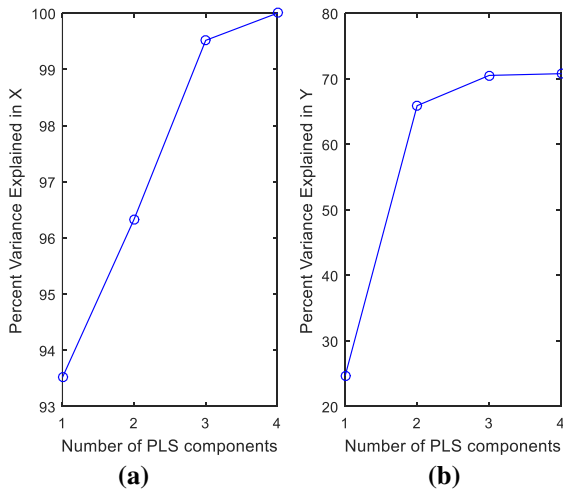
**Fig. 8.** Prediction models' performance overview.

remaining data then served as testing data and was used to validate the accuracy of the models. This testing set consisted of spectral data from 9 cakes of each sponge type. Each cake was of distinct age, and one cake of each sponge type was baked on the same day that hyperspectral data was captured. By including freshly produced cakes in the testing class, we aimed to determine if the PLS-R model performs adequately for unseen data.

When applying each model, the aim is to explain the variance in the data with 100% accuracy and to make a prediction with an error as close as possible to zero. By

observing the variance in predictors and responses and analysing the prediction error, it is possible to assess the performance of a model as a function of model complexity. While hundreds of wavelength bands were present in the hyperspectral data set, using them all to iteratively train the PLS-R model – which would result in many PLS-R components – can lead to strong overfitting of the model and redundant complexity for highly correlated data. In this study, an adequate number of PLS-R components was identified via 10-fold cross validation and observation of the mean squared prediction error of the trained models. Figure 6 shows the





**Fig. 9.** Graphs of variance explained in (a) predictors and (b) responses for chocolate cake PLS-R prediction model trained on multispectral data.

percentage of the variance described for the HSI data predictors and white cake moisture responses and the model prediction error when using an increasing number of PLS-R components. For the HSI PLS-R model, the use of 10 components (out of a maximum of 225) was deemed adequate; for the multispectral PLS-R model, 4 components (out of a maximum of 4) were used; and for the single band PLS-R model, the single component available was used.

Another way of assessing the prediction performance is by plotting the predicted values and comparing them with the ground-truth moisture and hardness measurements. An ideal model would have all the points lying on the diagonal of such a graph; this would mean that all predicted values are identical to those that have been observed. Figure 7 shows a comparison of the observed ground-truth measurements with predicted values for a PLS-R model trained and tested on HSI data.

A third method for assessing a PLS-R model's performance is to analyse its R-squared ( $R^2$ ) value. This is a statistical measure which assesses how closely the data points are to the fitted regression line. A value of  $R^2 = 1$  indicates a perfect fit, where the predicted values exactly match the observed measurements. Therefore, when training a PLS-R model, the aim is to get the  $R^2$  values as close as possible to 1.

#### 4. Results

Following training of the three models described in Sec. 3.4 for both types of sponge, each model was used to predict the moisture and hardness of unseen sponges of various ages. Table 1 compares the values of observed and predicted properties of the sponges – and the corresponding  $R^2$  values for the prediction of each characteristic – when a PLS-R model trained on HSI data was used.

Employing comparison plots and  $R^2$  values, Figure 8 shows an overview of the performance of all six models built from the acquired data. The figure shows the results applying three different models built from three different predictors' data sets (hyperspectral, multispectral and single band) for white sponges and for chocolate sponges. It can be seen in Figure 8 that the prediction of the hardness for chocolate cakes with multispectral data has slightly better performance than for a full hyperspectral profile (as demonstrated by  $R^2$

value). One explanation for this could be that the prediction model for these cakes was overfitted when using HSI data; this resulted in noise inadvertently being used to train the model. Further fine tuning of the model would be required before industrial implementation. Such tuning would consider the selection of the optimal spectral bands for each inspected cake type.

Using the multispectral imaging approach or single band camera has the potential to simplify the complexity of the hardware required to capture the data and reduce the associated monetary cost; however, it also introduces limitations to the accuracy and robustness of the prediction model. With a limited set of predictors, the variance in the responses may not be fully explained and, therefore, the performance of any resulting prediction model is limited. Figure 9 illustrates the variance explained in the predictors and responses for a PLS-R prediction model trained on multispectral data (simulated by the extraction of 4 bands from the HSI data cube); from Figure 9(b), it is clear that only around 70% of the variance in the responses can be described using the 4 selected bands.

A reduction in the spectral resolution of the predictors' data set clearly has an impact on PLS-R model performance; however, the accuracy of the prediction of moisture is significantly less affected by this limitation than the prediction of hardness. If a multispectral imaging system is preferred, optimisation of the band selection process would be necessary to select imaging wavelengths that better explain the variations in moisture and hardness.

#### 5. Conclusions

Hyperspectral imaging has been shown to be a valid method for predicting sponge cakes' moisture content and hardness. The application of this technology in the cake production environment, empowered by sophisticated image processing and prediction algorithms has huge potential. This research has established a proof of concept for a new stand-off cake moisture and hardness monitoring system. Such a system would provide the added advantage of being able to record every product in a hyperspectral image, which could be processed to detect variations in the production process.

The results of the moisture prediction when using a PLS-R model trained and tested on HSI data are very close to the observed values of the ground truth, with  $R^2$  values very close to 1. This proves that the developed technique provides a very accurate model for moisture prediction. Additionally, our study revealed that there is a degree of correlation between hardness and the moisture of the sponges; thus, performing hardness prediction based on reflectance data is justified and has shown promise. To the authors' knowledge, this has not been shown before.

The developed models appear to perform better for measuring moisture in white sponges than in chocolate sponges. The latter demonstrate a narrower range of moisture change over their lifetime and they exhibit a higher variance in overall sponge moisture, which leads to a weaker correlation between ground-truth measurements and predicted values for this type of sponge. Models based on hyperspectral data have the best moisture and hardness prediction performance; however, the other options (multispectral or single band) are still able to provide satisfactory results for sponge moisture.

## 6. Acknowledgements

This project was supported in part by the EPSRC Centre for Doctoral Training in Applied Photonics, funded by the UK Engineering and Physical Sciences Research Council (grant EP/G037523/1) and by an Industrial Fellowship from the Royal Commission for the Exhibition of 1851.

## 7. References

- [1] A. F. H. Goetz, G. Vane, J. E. Solomon, and B. N. Rock, "Imaging Spectrometry for Earth Remote Sensing," *Science*, vol. 228, pp. 1147-1153, June 7, 1985.
- [2] R. Hsuan and I. C. Chein, "Automatic spectral target recognition in hyperspectral imagery," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 39, pp. 1232-1249, 2003.
- [3] D. Bannon, "Hyperspectral imaging: Cubes and slices," *Nature Photonics*, vol. 3, pp. 627 - 629, 2009.
- [4] T. Kurz, S. Buckley and J. Howell, "Close-range hyperspectral imaging for geological field studies: workflow and methods," *International Journal of Remote Sensing*, vol. 34, no. 5, p. 1798-1822, 2013.
- [5] M. Denk, C. Gläßer, T. H. Kurz, S. J. Buckley and P. Drissen, "Mapping of iron and steelwork by-products using close range hyperspectral imaging: A case study in Thuringia, Germany," *European Journal of Remote Sensing*, vol. 48, no. 1, pp. 489-509, 2015.
- [6] G. Mirschel, O. Daikos, C. Steckert, K. Heymann and T. Scherzer, "Characterisation of sizes on textiles by in-line NIR chemical imaging," in *OCM 2017 - Optical Characterization of Materials*, Karlsruhe, 2017.
- [7] G. Mirschel, O. Daikos, T. Scherzer and C. Steckert, "Near-infrared hyperspectral imaging of lamination and finishing processes in textile technology," *NIR news*, vol. 28, no. 1, pp. 20 - 25, 2017.
- [8] C. Blanch-Perez-del-Notario and A. Lambrechts, "New Paper on Hyperspectral Imaging for Textile Recycling," 1 January 2016. [Online]. Available: [http://www.resyntex.eu/images/downloads/HSI\\_2016\\_paper\\_final.pdf](http://www.resyntex.eu/images/downloads/HSI_2016_paper_final.pdf). [Accessed 11 May 2017].
- [9] Y. Roggo, A. Edmond, P. Chalus and M. Ulmschneider, "Infrared hyperspectral imaging for qualitative analysis of pharmaceutical solid forms," *Analytica Chimica Acta*, vol. 535, pp. 79-87, 2005.
- [10] G. Mirschel, O. Daikos, T. Scherzer and C. Steckert, "Hyperspectral Imaging Used For In-Line Monitoring In Textile Technology," in *IASIM - 6th Conference in spectral imaging*, Chamonix-Mont-Blanc, France, 2016.
- [11] S. Marshall, T. Kelman, T. Qiao, P. Murray and J. Zabalza, "Hyperspectral Imaging For Food Applications," in *23rd European Signal Processing Conference (EUSIPCO)*, Nice, 2015.
- [12] D. Wu and D.-W. Sun, "Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review — Part II: Applications," *Innovative Food Science & Emerging Technologies*, vol. 19, pp. 15-28, 2013.
- [13] M. Dale, "Trash Talk," *Imaging & Machine Vision Europe*, pp. 20-22, December 2016.
- [14] Y. Lu, Y. Huang and R. Lu, "Innovative Hyperspectral Imaging-Based Techniques for Quality Evaluation of Fruits and Vegetables: A Review," *Applied Sciences*, vol. 7, no. 189, 2017.
- [15] D. S. Jayas, C. B. Singh and J. Paliwal, "Classification of wheat kernels using near-infrared reflectance hyperspectral imaging," in *Hyperspectral imaging for food quality analysis and control (1st ed.)*, San Diego, California, USA, Academic Press/Elsevier, 2010, p. 449-470.
- [16] H. Huang, L. Liu and M. O. Ngadi, "Recent Developments in Hyperspectral Imaging for Assessment of Food Quality and Safety," *Sensors*, vol. 14, pp. 7248-7276, 2014.
- [17] G. K. Naganathan, L. M. Grimes, J. Subbiah, C. R. Calkins, A. Samal and G. E. Meyer, "Visible/near-infrared hyperspectral imaging for beef tenderness prediction," *Computers and Electronics in Agriculture*, vol. 64, no. 2, pp. 225-233, 2008.
- [18] T. Qiao, J. Ren, C. Craigie, J. Zabalza, C. Maltin and S. Marshall, "Comparison between near infrared spectroscopy and hyperspectral imaging in predicting beef eating quality," in *Hyperspectral Imaging and Applications Conference (HSI 2014)*, Coventry, 2014.
- [19] T. Qiao, J. Ren, C. Craigie, J. Zabalza, C. Maltin and S. Marshall, "Quantitative prediction of beef quality using visible and NIR spectroscopy with large data samples under industry conditions," *Journal of Applied Spectroscopy*, vol. 82, no. 1, 2015.
- [20] T. Qiao, J. Ren, J. Zabalza and S. Marshall, "Prediction of lamb eating quality using hyperspectral imaging," in *OCM (Optical Characterization of Materials) 2015*, Karlsruhe, 2015.
- [21] M. S. Andresen, B. S. Dissing and H. Løje, "Quality assessment of butter cookies applying multispectral imaging," *Food Science & Nutrition*, vol. 1, no. 4, p. 315-323, 2013.
- [22] O. Bensaeed, A. Shariff, A. B. Mahmud, H. Shafri and M. Alfatni, "Oil palm fruit grading using a hyperspectral device and machine learning algorithm," in *7th IGRSM International Remote Sensing & GIS Conference and Exhibition*, Kuala Lumpur, 2014.
- [23] J. Li, X. Rao and Y. Ying, "Detection of common defects on oranges using hyperspectral reflectance imaging," *Computers and Electronics in Agriculture*, vol. 78, pp. 38-48, 2011.
- [24] R. Lu, "Detection of bruise on apples using near-infrared hyperspectral imaging," *Transactions of the ASAE*, 2003.
- [25] Y. Shuqin, H. Dongjian and N. Jifeng, "Predicting wheat kernels' protein content by near infrared hyperspectral imaging," *International Journal of Agricultural and Biological Engineering*, vol. 9, no. 2, p. 163-170, 2016.
- [26] P. Talensa, L. Morab, N. Morsyc, D. F. Barbind, G. ElMasryc and D.-W. Sund, "Prediction of water and protein contents and quality classification of Spanish cooked ham using NIR hyperspectral imaging," *Journal of Food Engineering*, vol. 117, no. 3, p. 272-280, 2013.

- [27] A. Apan, R. Kelly, S. Phinn, W. Strong, D. Lester, D. Butler and A. Robson, "Predicting Grain Protein Content in Wheat Using Hyperspectral Sensing of In-season Crop Canopies and Partial Least Squares Regression," *International Journal of Geoinformatics*, vol. 2, pp. 93-108, 2006.
- [28] M. Manley, "Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials," *Chemical Society Reviews*, vol. 43, pp. 8200-8214, 2014.
- [29] H. Kobori, N. Gorretta, G. Rabatel, V. Bellon-Maurel, G. Chaix, J.-M. Roger and S. Tsuchikawa, "Applicability of Vis-NIR hyperspectral imaging for monitoring wood moisture content (MC)," *Holzforschung*, vol. 67, no. 3, pp. 307-314, 2013.
- [30] S. Tsuchikawa and H. Kobori, "A review of recent application of near infrared spectroscopy to wood science and technology," *Journal of Wood Science*, vol. 61, no. 3, p. 213-220, 2015.
- [31] B. Leblon, "Overview on sensing technologies for real-time monitoring of wood properties," in *IASIM - 6th Conference in spectral imaging, Chamonix-Mont-Blanc, France, 2016*.
- [32] Y.-Y. Pu and D.-W. Sun, "Hyperspectral imaging in visualizing the non-uniform drying of mango slices during hot-air and microwave-vacuum drying," in *IASIM - 6th Conference in spectral imaging, Chamonix-Mont-Blanc, France, 2016*.
- [33] Z. Liu and F. Møller, "Bread Water Content Measurement Based on Hyperspectral Imaging," in *Scandinavian Workshop on Imaging Food Quality 2011, Ystad, 2011*.
- [34] M. B. Whitworth, S. J. Millar, and A. Chau, "Food quality assessment by NIR hyperspectral imaging," in *Proc. SPIE 7676, Sensing for Agriculture and Food Quality and Safety II, 2010*.
- [35] C. BRI. (2013). Bakery Product Development. Available: <http://www.campdenbri.co.uk/services/bakery-product-development.php>
- [36] M. W. Davidson, "ZEISS Microscopy Online Campus | Tungsten-Halogen Lamps," [Online]. Available: <http://zeiss-campus.magnet.fsu.edu/articles/lightsources/tungstenhalogen.html>. [Accessed 15 November 2018].
- [37] Labsphere, Inc, "Spectralon Diffuse Reflectance Standards," [Online]. Available: [https://www.labsphere.com/site/assets/files/1827/pb-13021-000\\_rev\\_02\\_og\\_spectralon.pdf](https://www.labsphere.com/site/assets/files/1827/pb-13021-000_rev_02_og_spectralon.pdf). [Accessed 15 November 2018].
- [38] BROOKFIELD ENGINEERING LABORATORIES, INC, "BROOKFIELD CT3 Texture Analyzer, Operating Instructions, Manual No. M08-372-E0315," [Online]. Available: <http://www.brookfieldengineering.com/download/files/CT3manual.pdf>. [Accessed 11 May 2017].
- [39] M. Kamruzzaman, Y. Makino and S. Oshita "Online monitoring of red meat color using hyperspectral imaging," *Meat Science*, vol. 116, p. 110-117.
- [40] J. L. Rodgers and W. A. Nicewander, "Thirteen Ways to Look at the Correlation Coefficient," *The American Statistician*, vol. 42, no. 1, pp. 59-66, 1988.
- [41] University of the West of England, "Pearson's Correlation Coefficient," [Online]. Available: <http://learntech.uwe.ac.uk/da/Default.aspx?pageid=1442>. [Accessed 11 May 2017].
- [42] P. Geladi and B.R.Kowalski, "Partial Least Squares Regression: A Tutorial," *Analytica Chimica Acta*, 185 (1986) 1-17.
- [43] S. Wold, M. Sjöström and L. Eriksson, "PLS regression: A basic tool of chemometrics," *Chemometrics and Intelligent Laboratory Systems*, 58, 109-130, 2001.
- [44] S. Wold, L. Eriksson, J. Trygg and N. Kettaneh, "The PLS method -- partial least squares projections to latent structures -- and its applications in industrial RDP (research, development, and production)" *Umea University, Umea, Sweden, 2004*.
- [45] H. Abdi, "Partial least square regression, Projection on latent structures Regression, PLS-Regression," *Wiley Interdisciplinary Reviews: Computational Statistics*, 2 2010, 97-106.
- [46] H Li, MA Lavin and RJ LeMaster, "Fast Hough Transform: a hierarchical approach," *Comput. Vision, Graphics & Image Process.*, Vol 36 (1986), pp. 139-161.