Rapid and Nondestructive Determination of Egg Freshness Category and Marked Date of Lay using Spectral Fingerprint

Thomas O. S. Akowuah,1,2 Ernest Teye,3 Julius Hagan,1 and Kwasi Nyandey4

1Department of Animal Science, University of Cape Coast, College of Agriculture and Natural Sciences, School of Agriculture, Cape Coast, Ghana
2Spark Design Company Limited, P. O. Box CT 7902, Accra, Ghana
3Department of Agricultural Engineering, University of Cape Coast, College of Agriculture and Natural Sciences, School of Agriculture, Cape Coast, Ghana
4Department of Physics, University of Cape Coast, College of Agriculture and Natural Sciences, School of Agriculture, Cape Coast, Ghana

Correspondence should be addressed to Ernest Teye; ernest.teye@ucc.edu.gh

Received 3 June 2020; Revised 20 September 2020; Accepted 4 October 2020; Published 15 October 2020

1. Introduction

Chicken egg is the most popular type of egg among the poultry family worldwide. It is consumed by a wide range of people globally as it cuts across people of different socio-economic strata. Egg is known to contain high-quality proteins, be rich in others such as carbohydrates, minerals, and has easily digestible fats [1]. The recent education on the health benefits of egg consumption has resulted in intensive production and consumption worldwide. However, one of the main concerns to consumers and quality control officers is the rapid and accurate determination of freshness. This is particularly important as increased consumption has resulted in many producers flooding the market with different categories of eggs. Particularly, this glut in supply has brought about storing eggs for longer periods before they are bought or sold, and this phenomenon leads to the reduction of freshness as a result of the influence of storage days on egg quality [2]. Some producers store eggs during the glut production and mislabel it as fresh (day-old eggs) for the market when there is high demand. In West Africa, the popular question mostly asked when buying egg is “how old is the egg” as consumers have related the storage duration to the freshness of the egg. This perception by consumers is supported by other researchers who found out that Haugh units is reduced with storage [2, 3]. Hence, many retailers are...
forced to sell their eggs at give-away prices or compelled to mislabel.

The quality of egg mostly perceived by consumers is its freshness, cleanliness, weight, shell quality, albumen index, Haugh unit, and chemical composition [2, 4]; however, the most popular method used to determine egg quality is Haugh unit (HU) developed by Haugh [5]. The freshness of an egg is the characteristic that is related to egg quality, and this attribute declines after laying [6, 7]. More so, during storage eggs become susceptible to internal quality deterioration; thus, there is a decline in quality as factors such as temperature, time, humidity, air movement, and handling are associated with deterioration [2]. Also, storage duration is noted to have a great deal of influence on egg freshness [8]; however, the marked date of lay on the shell of egg introduced by the European Commission regulation (2003/2295/EC) alone does not provide enough guarantee for egg quality [3]. Hence, a combination of factors is needed.

On the other hand, the analytical methods used for egg quality evaluation are often highly time-consuming, destructive, labour intensive, and often require sophisticated laboratory and cumbersome samples preparation. In this regard, various researchers have developed quality detection techniques such as electronic nose-based system [9], and this is based on the fact that fresh eggs have a very low concentration of organic volatiles, which increase during storage [10]. Another nondestructive operation consists of observing eggs against light and detecting the air cell. These methods cannot perform effectively during the first days after laying [3, 11]. Earlier studies involving a wider wavelength NIR spectroscopy have been used for the identification of freshness and quality assessment of eggs [12–15]. The application of portable NIR calibration model based on machine learning to determine egg storage time at room temperature has also been done [16]. These studies did not consider the simultaneous measurement of freshness eggs under two conditions to cover peculiar situations in developing countries (often challenged with cold storage infrastructure). More so, little or no work has been done using portable NIR spectroscopy for simultaneous classification and prediction of the freshness of egg under different storage conditions to represent marked date of lay. This will facilitate the rapid quality control and checking of fraudulent mislabelling of the marked date of lay on the shell as introduced by European regulations.

NIR spectroscopy is more convenient, nondestructive, rapid, and simple analytical techniques, which require little or no elaborate sample preparation [17]. Furthermore, the advances in computer and electronics that have resulted in miniaturizing NIR spectrometers have added advantage by making this technology applicable and user-friendly outside the laboratory. For NIR spectroscopy, the relative contribution of reflected and absorbed radiation depends on the chemical composition, microstructure, and physical parameter of the material under consideration [17, 18]. This current study, which is the first in Ghana and West Africa, aims at developing a novel rapid nondestructive and simultaneous detection of egg freshness and marked date of lay of eggs in two different storage conditions by using a portable NIR spectrometer coupled with multivariate algorithms. This means that our technique will importantly check egg fraud due to mislabelling in Africa. It will also lead to the comprehensive quality control and quality assurance monitoring in the egg value chain. The novelty of these studies lies in the simultaneous detection of intact egg freshness and storage days in both ambient and cold storage conditions. In this regard, egg production and sale will be monitored and controlled irrespective of the storage type used to reduce food fraud.

2. Materials and Methods

2.1. Egg Samples. In this study, 120 fresh red intact eggs were sampled from Lohman Brown layer strain aged between 36 and 64 weeks directly from the School of Agriculture teaching and research farm of the University of Cape Coast in the Central region of Ghana and transported to the research laboratory for further analysis. The samples were divided into two: one part that is 60 eggs were stored in the cold storage of 4°C, using tabletop fridge, while the other (60 eggs) was stored in ambient temperature (28°C) with a relative humidity of 70%. Furthermore, in each storage condition, 30 eggs were separately used to determine Haugh unit and York height that were determined as a measure of freshness during storage by employing the recommended method. For every five days, five eggs were randomly selected from each storage type for these measurements.

2.2. Sample Spectra Acquisition. The spectrum of each egg was scanned continuously from the day of lay (fresh sample collection day) to twenty days (0–20 days) in the reflectance mode using a handheld spectrometer (SCIO™) with spectra range of 740 nm and 1070 nm in a 1 nm resolution for spectra data recording. For each egg, the equatorial region of the eggshell was scanned three times after rotating it at 120°. This portion was selected because the internal composition changes are mostly more significant in that region compared to the others [19]. The scanning was done at an ambient temperature of 28.6 ± 1°C with a humidity of 68%.

2.3. Reference Measurements of Freshness Using Haugh Units. The Haugh unit (HU) of the eggs (as a reference method for freshness) was computed by the method used (with equation (1)), while the York height was also measured using a digital Vernier caliper according to the method used by other authors [20]. To monitor the freshness of the eggs using HU, ten (10) eggs from each storage group were used to determine the Haugh unit and York height during the entire storage period from day zero to twenty days (0–20 day’s storage period).

\[
HU = 100\log\left(h + 7.6 - 1.7w^{0.37}\right),
\]  

where HU = Haugh unit, \(h\) = albumen height (mm) by using a digital Vernier caliper and \(w\) = egg weight (g) by using digital weighing scale (0.001 g).
2.4. Software Device. Spectra data recordings stored in a cloud-based data-set with their corresponding reference values for the time of scanning were downloaded using a research license of SCIO lab and imported to MATLAB version 9.5.0 (Mathworks Inc., USA) with windows 10 Basic for data processing for all preprocessing treatments and multivariate algorithms.

2.5. Data Partition. The raw data were divided into two subsets, calibration set for developing the model and prediction set for evaluating the predictive ability of the constructed models. To avoid bias, 75% of data from the samples were selected as the calibration set, while the remaining data were the prediction set. To achieve so, the members in each set were selected to come to a 3/1 division of calibration set/prediction set.

2.6. Data Preprocessing. On the spectra, the raw data-set was preprocessed with multiplicative scatter correction (MSC) because the models developed using the raw spectra data usually do not give the desired results. MSC tool is a useful technique for the correction of scattered light and inclination of baseline variation [21, 22]. Preprocessing the spectra data is an integral part of modelling to eliminate background information and noise from the useful properties of the scanned samples [16, 23]. Principal component analysis (PCA) was done to observe any known cluster trends. PCA is an unsupervised data description and dimension reduction techniques mostly used to perform cluster analysis in spectra data [24]. This is normally done before any multivariate modelling to detect patterns from the data matrix as it brings out visualized data trends in dimensional space [25].

2.7. Multivariate Analysis Methods. LDA as a linear parametric classification technique was employed in this work. It performs its function by maximizing the between-class variance over the within-class variance to create a linear decision boundary between them and find linear combination of features that best differentiates two or more groups of events [26]. The principle of LDA is based on the determination of linear discrimination functions and the number of principal component factors is crucial to its performance [27]. In this study, PCA data was used as an input data for the LDA to build the identification model.

Partial least squares regression (PLS-R) on the other hand was used for predicting the marked date of lay. PLS-R is a popular linear multivariate tool that analyzes data with strong collinear, noisy, and redundant variables; for more information, refer to other authors [28, 29]. The performance of LDA was evaluated by identification rate (%), while PLS model was evaluated by using three main parameters, namely, the root mean square error of cross-validation (RMSECV), the root mean square error of prediction (RMSEP), and the correlation coefficient (R) [30]. These parameters were calculated by equations (2)–(5):

\[
IR = \frac{n_1}{n} \times 100, \quad (2)
\]

\[
RMSECV = \sqrt{\frac{\sum_{i=1}^{n}(\tilde{y}_i - y_i)^2}{n}}, \quad (3)
\]

\[
RMSEP = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \tilde{y})^2}{n}}, \quad (4)
\]

\[
R = \sqrt{1 - \frac{\sum_{i=1}^{n}(\tilde{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \tilde{y})^2}}, \quad (5)
\]

where \(n_1\) = number of samples correctly identified, \(n\) = the number of samples, \(y_i\) = the reference measurement results for sample \(i\), \(\tilde{y}_i\) = the estimated result for sample \(i\) when the model is constructed with sample \(i\) removed, \(\tilde{y}\) = the mean of the reference measurement results for all samples.

3. Results and Discussion

3.1. Egg Freshness Measurements. To monitor and confirm the freshness of egg under two storage conditions, destructive standard method was used. As seen from Table 1, it was observed that egg freshness decreases or is reduced with storage duration using the Haugh unit (HU) rating. This observed phenomenon is in agreement with finding reported by other authors [8]. Within the HU rating scales, 72 and above are graded as AA, 60–72 as A, and lower than 60 as B, while below 60 is also known as consumer resistance point and below 50 is considered poor and unacceptable, respectively [12]. However, for eggs stored in cold storage condition, HU values do not correspond directly to freshness or duration in storage. This information revealed, therefore, that classifying and predicting the storage duration of eggs under well-defined storage category provide useful information to consumers. Furthermore, this table also supports the idea that developing a model to predict egg quality under two storage conditions is very vital. It was observed that eggs stored in the cold storage condition were still very good with regards to HU even at 20 days of storage duration.

| Table 1: Reference measurement egg quality under two storage conditions. |
|-----------------------------|-----------------------------|
| Days | Cold storage Haugh unit | Cold storage York height | Ambient storage Haugh unit | Ambient storage York height |
| 0    | 96.78 | 19.3 | 95.06 | 18.7 |
| 1    | 95.56 | 18.6 | 87.12 | 18.2 |
| 5    | 88.97 | 18.4 | 77.20 | 16.9 |
| 10   | 86.48 | 18.3 | 71.89 | 13.8 |
| 15   | 82.53 | 15.9 | 60.01 | 11.8 |
| 20   | 82.42 | 18.2 | 41.36 | 9.9 |
3.2. Spectra Presentation. The spectral profile shown in Figures 1(a) and 1(b) revealed that it is difficult to differentiate between the spectra of eggs stored at different days for under ambient storage and cold storage. It could be seen also from Figure 2 that the differences observed were as a result of the mean plot for the eggs stored from 0 to 20 days. These figures revealed both useful and redundant information. Furthermore, the eggs were grouped into four classes plus a zero-day class. The mean plot as seen in Figure 3 revealed a distinct spectral profile. It could be explained that each category of eggs showed a characteristic fingerprint of CH 3rd overtone and NH 2nd overtone region that made it distinguish itself. Furthermore, a PCA was used as unsupervised pattern recognition techniques to observe a well clustered trend.

PCA was used as an unsupervised tool to identify cluster trend in the spectra data. From the results obtained it was observed that PCA after preprocessing with MSC gave some clear separations as observed for PCA results at ambient storage and cold storage as seen in Figure 4. The total 3 PCs were 99.95% for ambient storage and 99.91% for cold storage. These mean that the three main PCs contributed to the clear observed cluster trend when MSC was used on the raw data.

Also the PCA for the four classes studied (Figure 5) further revealed that raw spectra did not give any separation, while MSC-PCA gave a neat separation as seen in Figures 5(a) and 5(d). It was, however, observed that cold storage data separated more neatly than the ambient storage condition. This could be a result of a more stable storage condition in the cold storage compared with the ambient storage.

3.3. Classification Models. LDA algorithms were used for developing an identification model for identifying egg freshness in the four categories based on storage durations. The result of the classification models seen in Table 2 reveals that MSC-LDA performed better with classification rate above 96% in both calibration set and prediction set at optimal principal components of 5 for eggs stored in ambient and cold conditions, respectively.
3.4. Quantification Model. In this research, another model was simultaneously developed to predict the specific egg storage duration under ambient and cold storage conditions. PLS-R model was used for predicting these storage durations under both storage conditions. From Figures 6(a) and 6(b), it could be seen that the measured values correlated well with NIR predicted values for both storage conditions. This good relationship was confirmed in Table 3 with $R$ above 86%. The measured values correlated linearly with NIR predicted measurements. However, there were a few outliers that subsequently affected the PLS model. From Table 3, it could be seen that the MSC-PLS model was the best with...
parameters of $R$ above $= 0.85$ and RMSEC under 3.3 days for both storage conditions in the calibration set and prediction set, respectively.

3.5. General Discussion. The application of a portable spectrometer for fingerprinting egg samples resulted in creating a spectral profile, which was unique to the freshness composition of the samples used. The fingerprinting wavelength range of 740 nm and 1070 nm provided some information (chemical and physical properties) that could be vital for classifying and predicting egg freshness categories and marked date of lay. The spectral profile shows multiply bands and some peaks as seen from Figure 7. These bands are made up of overtones and combinations of fundamental vibrations, which correspond to useful chemical and physical properties in the categories of eggs used and their corresponding freshness. These properties could be useful for qualitative and quantitative fingerprints, because spectrum is changed due to physical and chemical interactions when it is passed through a material and thus can be compared with the changed spectrum, and optical information details of such biological material could be linked with the chemical and physical quality [31]. Hence, there was the need to use advanced mathematical models through preprocessing to extract these vital information. MSC preprocessing technique provided the need pretreatment for the spectral data-set. From the results obtained, it was obvious that MSC improved the performance of both the classification and the quantification model in this study. It means that MSC techniques performed the correction of scattered light and inclination of baseline variation well as proposed by other researchers [21, 22]. It also supports the belief that preprocessing spectra data is an integral part of modelling to eliminate background information and noise from the useful properties of the scanned samples [16, 23]. Furthermore, the optimum classification results obtained revealed that there was a linear correlation between the NIR spectral and the egg freshness categories studied. For the determination of the marked date of lay, PLS-R was used (for the determination of storage duration). From the results obtained, the measured values (by using destructive techniques) correlated linearly with NIR predicted measurements. However, there were a few outliers that subsequently affected the PLS model. Figure 7(b) explains how the complexity of the PLS-R model was developed for predicting freshness in terms of storage duration. It also refers to how the fingerprint was considered by the model components and how to interpret the meaning.
of each model component [32]. The PLS-R weight plot for the first two PCs shows that the major peaks that contributed to its performance were around the wavelength of 840–855 nm, 875 nm, and 1000–1033 nm. These wavelengths correspond to OH second overtone, NH second overtone, and CH third overtone, which could be associated with pH, protein, and carbohydrates in biological materials. More importantly, the NH, CH, and OH overtones represent aromatic amino acids, which are important organic compounds in egg as it contains amine (NH₂) and carboxylic acid (-COOH) functional groups [33]. These functional groups are important components of proteins because proteins are made up of hundreds or thousands of smaller units of amino acids [33]. More so, research has shown that

![Figure 5: Raw (A1-B1) and MSC (A2-B2)-PCA score plot of eggs stored under ambient and cold storage.](image)

<table>
<thead>
<tr>
<th>Table 2: The overall performance of multivariate classification methods.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model</strong></td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Ambient storage</strong></td>
</tr>
<tr>
<td>LDA</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Cold storage</strong></td>
</tr>
<tr>
<td>LDA</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Figure 6: Measured versus NIR prediction (a) for ambient storage and (b) cold storage.

Table 3: The overall performance of PLS regression model.

<table>
<thead>
<tr>
<th>Pretreatment Factors</th>
<th>Calibration R</th>
<th>RMSEC (Days)</th>
<th>Prediction R</th>
<th>RMSEP (Days)</th>
<th>Independent R</th>
<th>RMSEI (Days)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ambient storage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw</td>
<td>9</td>
<td>0.61</td>
<td>0.59</td>
<td>6.47</td>
<td>0.61</td>
<td>6.85</td>
</tr>
<tr>
<td>MSC</td>
<td>5</td>
<td>0.83</td>
<td>0.89</td>
<td>3.12</td>
<td>0.87</td>
<td>2.57</td>
</tr>
<tr>
<td><strong>Cold storage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw</td>
<td>9</td>
<td>0.54</td>
<td>0.49</td>
<td>5.78</td>
<td>0.51</td>
<td>6.05</td>
</tr>
<tr>
<td>MSC</td>
<td>7</td>
<td>0.86</td>
<td>0.91</td>
<td>2.481</td>
<td>0.885</td>
<td>2.66</td>
</tr>
</tbody>
</table>

Figure 7: Continued.
the chemical and physical properties of egg change with storage time, temperature, and humidity [15, 34]; hence, these changes also contributed to the identification and quantification of egg freshness by the portable NIR spectroscopy technique. Specifically, the albumen pH of freshly laid eggs ranges from 7.6–8.5 and changes to 9.5 in the first few days resulting in the gel-like thick albumen [34, 35]. Therefore, the CH third overtone identified around 875 nm could be associated with pH in eggs. In addition, decomposition process in egg may be the cause of changes observed in the overtones for the eggs [36]. These findings agree with other findings reported by others [16]. The short OH overtone may be attributed to eggshell surface being moist shortly after the egg is laid, and it could be explained that moisture loss due to drying of the cuticle of the egg may cause gradual increment in the intensity of the OH band through time and simultaneously, cuticle becomes thinner and the eggshell carbonate mineral becomes more exposed to the surface [16].

Generally, the findings agree with other studies. Specifically, the R of 0.87 and 0.88 obtained for independent set for ambient and cold storage, respectively, were consistent with the findings reported by Coronel-Reyes and co-workers [15, 16, 37]; however, the RMSEC of 2.57 and 2.66 varied from these authors. For instance, Sun and others [37] obtained R value of 0.8653 and RMSECV of 3.745 for using artificial vision and dynamic weighing to assess egg freshness. Also, others obtained R value of 0.89 and RMSECV of 1.65 by using lab grade VIS/NIR spectroradiometer [15]. Furthermore, other studies by Aboonajmi and Abbasian Najafabadi [38] by using VIS/NIR spectral measurements from 300–1100 nm found a square regression of 0.79 for Haugh. The variation observed in these findings could be attributed to the differences in the predictive models and the development of new experimental setup [13].

4. Conclusion

The study has revealed that portable NIR spectroscopic techniques could be used for rapid nondestructive method for simultaneous analysis of eggs: for classification of egg freshness category and prediction of the marked date of lay. For the classification challenge, the PCA-MSC-LDA gave IR =< 95% in calibration set and prediction set for eggs stored under ambient and cold storage conditions. For predicting the storage duration, MSC-PLSR had a prediction performance of R = 0.83 and above for all the storage conditions investigated. This finding could be useful for supporting the utilization of handheld NIR spectroscopy for simultaneous determination of egg freshness categories and mark date of lay of eggs stored in either cold or ambient storage condition.

Data Availability

Data will be made available on request.

Conflicts of Interest

The authors, the authors declare that there are no conflicts of interest.

Acknowledgments

The authors highly appreciate the financial support provided by University of Cape Coast. The fresh Egg supplied by
School of Agriculture poultry farm is highly acknowledged. The authors also thank Francis P. Lamptey, Yazeed Pangdorou, for their support. The authors appreciate proof reading assistance provided by Mrs. Winifred Akpene Teye.

References


