Electronic blood detection in eggs: B. De Ketelaere et al.

Improved blood detection in consumption eggs using combined reflection-transmission spectroscopy

B. DE KETELAERE*, K. MERTENS, B. KEMPS, F. BAMELIS, B. KAMERS, E. DECUYPERE and J. DE BAERDEMAEKER

K.U. Leuven, Department of Agro-Engineering and -Economics, Egg quality and incubation research group; Kasteelpark Arenberg 30, 3001 Leuven, Belgium

*bart.deketelaere@biw.kuleuven.be

Keywords: egg; blood detection; spectroscopy; transmission; reflection

Summary
The presence of blood in consumption eggs is strongly rejected by consumers and egg grading companies address much attention to the topic. Grading eggs for the presence of blood is commercially applied using visual and near infrared transmission spectroscopy and yields acceptable results for white eggs but encounters difficulties in brown shelled eggs due to the presence of protoporphyrin in the shell, a molecule that resembles haemoglobin, the optically active component in blood. In this contribution, an improvement to the existing methodology is proposed that uses combined reflection and transmission measurements. The system has been validated on white and brown shelled eggs which have been injected with 50 µl venous chicken blood. It is shown that the methodology proves to be superior to transmission-only measurements. However, for brown shelled eggs the detection rate remains low, and further research is needed.

Introduction
Egg quality encompasses a number of aspects, related to the shell, the albumen and the yolk and may be divided into external and internal quality. The appearance of the egg (external quality), as influenced by the severity of defects, is important for consumer appeal. Egg shells are evaluated on the basis of cleanliness, shape, texture, and soundness. The internal quality is based on air chamber size, albumen quality, yolk quality, and the presence of blood or meat spots. The albumen is a major indicator of the overall interior egg quality. Thinning of the albumen is a sign of quality loss. When a fresh egg is carefully broken onto a smooth flat surface, the round yolk is in a central position surrounded by thick albumen. When a stale egg is broken, the yolk is flattened and often displaced to one side and the surrounding thick albumen has become thinner, resulting in a large area of albumen collapsed and flattened to produce a wide arc of liquid. This is the principle that is used in measuring Haugh units, and it is still commonly used to judge albumen quality (Haugh, 1937). Yolk quality is related to its appearance, texture, firmness, and smell.

A number of these quality aspects are routinely checked in packing houses through candling. Candling has the advantage of being non-destructive and rapid. Accurate candling is best done in a darkened room by passing light through each egg. Candling equipment may range from a simple homemade unit to a mechanical device which is part of a mechanized washing, grading, sizing, and packing unit. Regardless of the type of equipment used, each egg must be examined. However, the increasing throughput of modern egg sorting machines, which can grade up to 120 000 eggs per hour, makes this quality inspection step become more difficult and, hence, less reliable. For other quality aspects, such as eggshell strength or albumen quality, destructive and non-destructive techniques were developed years ago. Most of them have the disadvantage of being time consuming so that they can only be applied on a small sample of the considered batch (De Ketelaere et al., 2004).

The investigation of the possible use of optical spectroscopy for grading consumption eggs was initiated more than half a century ago when Norris and Brant found time related changes in the spectral data around 750 nm several hours after laying (unpublished results, quoted in Norris, 1996). However, the researchers did not find a relationship between these changes and the internal egg quality. One explanation could be that the change in optical scattering properties of the eggshell was measured, as the moisture content of the shell stabilizes during the first hours after laying.
More recently, Schmilovitch et al. (2002) showed that pH, being an important albumen quality parameter, could be determined using a Partial Least Squares (PLS) regression model on the first derivative of the NIR spectral data. Also Kemps et al. (2005a and b) showed that the VIS-NIR spectral data of intact eggs contain information about the albumen quality, in terms of the widely used Haugh units. A correlation of 0.82 was reported between the measured and predicted Haugh units, which is sufficient to sort eggs into different classes. In an attempt to make the technology low-cost, they showed that the Haugh units could also be predicted using only three different optical transmission ratios ranging from the visible to near infrared region of the spectrum. This can be done in a single measurement. In this way, a correlation of 0.7 was obtained between the measured and the predicted Haugh units.

Another important application of VIS-NIR spectroscopy for grading eggs is checking for blood inside the egg. Although the prevalence of blood containing eggs is low (< 1 %), it attracts much attention of egg grading companies because they are strongly rejected by consumers who often relate it to fertile eggs. When measuring transmission spectra of eggs, light travels through the egg and the absence or presence of internal optically active anomalies can be detected. The detection of blood inside eggs is already used in grading machines. The optically active pigment of blood, haemoglobin, has three main absorption peaks, namely at 415, 541 and 577 nm. As the calciferous shell of the egg absorbs all transmitted light beneath 550 nm, only the 577 nm absorption peak can be used to detect the presence of blood in eggs (Brant et al., 1953). In order to correct for eggshell thickness, eggs size and other non haemoglobin–related characteristics, the transmission at 577 nm is divided by the transmission at a reference wavelength. Depending on the author, a wavelength between 585 and 610 nm (Gielen et al., 1979) is chosen. The ratio between the two transmission values is called the ‘blood value’ and is used as classification criterion.

The detection success of blood spots in eggs is highly dependent on shell colour: while on white eggs high detection accuracy can be achieved in general, this is not the case for brown coloured shells. Indeed, one brown pigment of the eggshell, protoporphyrin, has optical properties that are closely related to haemoglobin: it has an adsorption peak at 589 nm, very close to the absorption peak of haemoglobin (577 nm) and this makes the detection of blood in brown shelled eggs particularly difficult. Even the use of a reference wavelength cannot solve this problem (Gielen et al., 1979). Some detectors can adapt the selection threshold based on a colour measurement of the eggshell, but the detection still remains difficult and needs further research. Own research (unpublished results) have even proved that also for white shelled eggs problems can occur because the human perception of shell colour is not a good measure of the protoporphyrin content. Indeed, for some genetic laying lines of white eggs, the protoporphyrin content of the shell is elevated resulting in grading difficulties.

Therefore, the objective of this research is to improve the blood detection in eggs by using a correction factor that is based on the combined reflection-transmission data of the eggs.

Materials and methods

MEASUREMENT SETUP
A labscale test set-up was constructed, using two identical spectrophotometers (Ocean Optics S2000). The light is sent through a 200nm slit onto a fixed grating. The grating of the spectrophotometer is optimised for the 200-900 nm bandwidth. A 2048 pixels CCD-array is used for detection. The opening of the shutter was regulated in such a way that the intensity of the light didn’t cause an overload in the spectrophotometer. One spectrophotometer was used for reflection measurements in order to correct for the protoporphyrin content, the other was used for transmission measurements. For reflection, one sensor was placed above the egg with angle of 45°. A second sensor was placed underneath the egg, at 180° from the light source. A software program was written in LabVIEW to simultaneously capture both signals, using a USB and serial port of a personal computer. A halogen lamp (150 Watt) was used as light source.

During the measurements, the egg was placed on two diabolo shaped rollers and spins around its long axis. For each egg, the average of three measurements along the equator of the egg was used throughout further analyses.

EGG MATERIAL AND BLOOD INJECTION
300 intact, fresh, white shelled eggs were used in a first experiment. At day 0, 114 eggs were injected with blood as follows. Venous chicken blood was collected from the wing of a hen, and collected into a
A small hole was drilled on the side of the air chamber of the egg and fresh blood (50 µl) was injected in the albumen using a fine syringe. After 4 days of storage, the 300 eggs were measured using the above-described set up, and were broken in order to investigate the blood injection. A COMBUR® test was used in order to investigate whether the blood had diffused into the albumen. Subsequently, a digital camera was used to take a snapshot of the blood injected eggs. 90 control eggs were used to calibrate the model; the remaining eggs (96 control and 114 blood containing) were used to validate the methodology.

In the same way a second experiment was conducted on 258 fresh, brown shelled eggs. 100 eggs were used to calibrate the model without injecting blood, 100 control eggs and 58 blood injected eggs served for validation of the models.

SPECTRAL PRE-PROCESSING AND STATISTICAL ANALYSIS

The collected spectral data (transmission and reflection) were first subjected to a smoothing procedure in which a Savitsky-Golay algorithm was used that fits a second degree polynomial to an 11 point spectral region. Afterwards, each spectrum was divided by the transmission, respectively reflection, at 610 nm as a normalization step.

The pre-processed spectral data were used in order to build a partial least squares (PLS) regression model that predicts the transmission at 577 nm (the characteristic wavelength of haemoglobin) using the whole reflection spectrum. Calibration was performed on control (non blood containing) eggs only. In order to obtain an optimal calibration model, a cross-validation procedure was used with repeated random test sets. For this purpose, the data were split into 5 sections and 10 iterations. Since the transmission data were converted by using only the value at 610 nm, it would allow for using only two photodiodes instead of a spectrophotometer, making the total cost of the system much more affordable.

For validation, the predictive calibration model was applied to the validation data, comprising both control and blood injected eggs. In order to judge the potential of the developed methodology, it was compared to the state-of-the-art methodology, which is solely based on transmission and is given by the ratio T577 nm / T610 nm. The MATLAB software was used throughout all analyses (The Mathworks, Inc., USA).

Results and discussion

WHITE SHELLED EGGS

Figure 1 shows the pre-processed calibration data (reflection and transmission) between 400 and 800 nm. It is clear from that figure that the different contents in pigments in the shell and other matter cause a very large variance in the measured spectral data.
Using the normalized spectral data, a calibration model was built as described in the materials and methods section. A plot of the RMSECV (Root Mean Squared Error of Cross Validation) values as a function of the number of latent variables indicated an optimal model complexity of 5 latent variables (Figure 2).

![Figure 2](image2.png)

**Figure 2** The RMSEC and RMSECV as a function of model complexity. 5 latent variables were chosen based on this information.

The model was capable of predicting the transmission at 577 nm with high accuracy ($r = 0.92$), using only 5 latent variables. The regression coefficients of this PLS model are given in Figure 3. It is clear that the protoporphyrin content of the shell, which has specific absorption bands at 485, 589 and 643 nm has a negative effect on the transmission at 577 nm. This observation points exactly to the difficulties seen in practice, namely the large biological variability in protoporphyrin content causing blood detection problems. When validating the calibration model on the control eggs in the test set, this correlation remained stable ($r = 0.91$) indicating that the model had a reasonable generic character and that no over fitting was present.

![Figure 3](image3.png)

**Figure 3** PLS regression coefficients using 5 latent variables. The influence of protoporphyrin, with characteristic absorption at 485, 589 and 643 nm) on the transmission at 577 nm (dependent variable) can readily be observed.
Subsequently, the actual values of the transmission at 577 nm were subtracted from predicted values of the transmission at 577 nm. We will call this quantity the prediction difference. For the calibration data, this prediction difference is zero on average, as expected. Since in practice it is mandatory to keep the number of false rejects (i.e. the downgrading of control eggs) to an acceptable level, the classification threshold will be based on the 97.5 % upper limit of the values of the prediction difference. This threshold is given by the average value for the ratio augmented by 1.96 times the standard deviation, and this for the 90 calibration eggs. As such the threshold for blood detection becomes $0 + 1.96 \times 0.0183 = 0.0359$.

The calibration model was subsequently used in order to estimate the transmission at 577 nm for the test set of eggs (96 control eggs and 114 blood injected eggs). It is expected that for the control eggs, the estimated and actual values are close to one other. However, in the case blood is present in the egg, the predicted value, which is based on the reflection spectrum and hence does not contain any information on blood, would be substantively higher than the actual value, due to the absorption of blood. Using the threshold on the validation set, there was one false reject (1.0 %) and 5 blood containing eggs were falsely classified as control (4.4 %). When linking these results to the COMBUR® data and the digital snapshots, it was clear that the false negative eggs were mainly eggs in which the blood was not diffused into the albumen, but was only present as one small, dense spot.

When using the classical approach that is only based on transmission, more specifically based on the ratio between the transmission at 577 nm (characteristic absorption of haemoglobin), and at 610 nm (acting as a reference wavelength) results were inferior. Also here a 97.5 % limit of the values of the ratio $T_{577 nm} / T_{610 nm}$ was constructed. However, since for blood containing eggs this ratio is lower than for control eggs due to the absorption of haemoglobin at 577 nm, this threshold is given by the average value for the ratio decreased by 1.96 times the standard deviation, and this for the 90 calibration eggs. As such the threshold for blood detection becomes $0.8053 - 1.96 \times 0.0442 = 0.7187$.

Applying this ratio to the validation set of control and blood containing eggs, 3 eggs are falsely rejected (3.1 %), and 12 blood containing eggs were classified as control eggs (10.5 %). The COMBUR® test revealed that also here, in the majority of the cases, the blood had not diffused into the albumen.

**BROWN SHELLED EGGS**

Figure 4 shows the pre-processed calibration data (reflection and transmission) between 400 and 800 nm. It is clear from that figure that the different contents in pigments in the shell and other matter cause a very large variance in the measured spectral data. The absorption bands of protoporphrin (at 589 and 643 nm) are clearly visible.

![Figure 4](image-url) Overview of the data for brown shelled eggs. Normalized reflection (left) and transmission (right).
Using the normalized spectral data, a calibration model was built as described in the materials and methods section. A plot of the RMSECV values as a function of the number of latent variables indicated an optimal model complexity of 7 latent variables (Figure 5).

![Figure 5](image)

**Figure 5** The RMSEC and RMSECV as a function of model complexity. 7 latent variables were chosen based on this information.

The model was capable of predicting the transmission at 577 nm with moderate accuracy ($r = 0.84$), using 7 latent variables. The regression coefficients of this PLS model are given in Figure 6. The protoporphyrin content of the shell, which has specific absorption bands at 485, 589 and 643 nm, seems to dominate the regression coefficients also here, which points again to the difficulties seen in practice. When validating the calibration model on the control eggs of the test set, this correlation remained stable ($r = 0.83$) indicating that the model had a reasonable generic character and that no over fitting was present.

![Figure 6](image)

**Figure 6** PLS regression coefficients using 7 latent variables. The influence of protoporphyrin, with characteristic absorption at 485, 589 and 643 nm) on the transmission at 577 nm (dependent variable) can readily be observed.

In analogy with the first experiment, the actual values of the transmission at 577 nm were subtracted from predicted values of the transmission at 577 nm (the prediction difference) and a classification
threshold based on the 97.5 % upper limit was constructed. The threshold for blood detection in brown eggs becomes \(0 + 1.96 \times 0.0265 = 0.0530\). The calibration model was used in order to estimate the transmission at 577 nm for the test set of eggs (100 control eggs and 58 blood injected eggs). Using the threshold on the validation set, there were two false rejects (2.0 %). On the contrary, only 18 of the 58 blood eggs could be detected in this way, 31 %. In the case that blood had diffused into the albumen, the detection rate was acceptable (8 / 10 eggs, 80%). However, if no diffusion had occurred and blood was present as one dense spot, the detection rate was mediocre (10 / 48 = 21%).

When using the classical approach that is only based on transmission, more specifically based on the ratio between the transmission at 577 nm (characteristic absorption of haemoglobin), and at 610 nm (acting as a reference wavelength) results were inferior. Also here a 97.5 % limit of the values of the ratio \(T_{577\text{ nm}} / T_{610\text{ nm}}\) was constructed. However, since for blood containing eggs this ratio is lower than for control eggs due to the absorption of haemoglobin at 577 nm, this threshold is given by the average value for the ratio decreased by 1.96 times the standard deviation, and this for the 90 calibration eggs. As such the threshold for blood detection becomes \(0.2100 - 1.96 \times 0.0539 = 0.1044\).

Applying this ratio to the validation set of control and blood containing eggs, 3 eggs are falsely rejected (3.0 %), and only 3 blood containing eggs were detected (5.2 %). The COMBUR® test revealed that in the three detected blood eggs the blood had diffused into the albumen.

**Conclusions**

The results presented in this paper clearly show that the problem of blood detection in brown eggs is far from solved. However, the research show that the proposed methodology proves to be an important enhancement of the current methodology based on transmission only spectral data. Furthermore, by using a data pre-processing which is mainly based on dividing each light intensity by the intensity at 610 nm, the technique would allow for using only two photo cells for transmission data instead of a broad range spectrophotometer. This would lower the cost of implementation by a substantial amount.

The very low detection rates also raise the question whether the proposed blood injection methodology is mimicking the natural blood eggs in a reliable way since commercial egg grading machinery manufacturers claim higher detection rates than those reported in this study. The main difficulty in this type of research, however, is that obtaining natural blood eggs is extremely difficult for two main reasons. First, the prevalence of blood in eggs is low (in the order of 1 %) so that a very large amount of randomly selected eggs would be needed in order to obtain a sufficient amount of blood eggs. Second, there is as of this moment no non-destructive detection principle available that can detect all blood eggs. Taking rejected eggs from a commercial egg grader is no option since it would result in a biased dataset that probably only contains easy-to-detect blood containing eggs.

Probably the only way to perform further research on this methodology and to get a clear view on the detection rates is to convert the proposed methodology into an on-line system that can be integrated into a commercial grading line, and to inspect all eggs that were rejected or accepted by the methodology.

**Acknowledgements**

This research has been funded by the European Project CT-2001-01606 “EggDefence”.

**References**