Monitoring Quality Loss of Pasteurized Skim Milk Using Visible and Short Wavelength Near Infrared (SW-NIR) Spectroscopy (600-1100 nm) and Multivariate Analysis

Hamzah M. Al-Qadiri\textsuperscript{a*}, Mengshi Lin\textsuperscript{b}, Murad A. Al-Holy\textsuperscript{c}, Anna G. Cavinato\textsuperscript{d}, and Barbara A. Rasco\textsuperscript{e}

\textsuperscript{a}Department of Nutrition and Food Technology, Faculty of Agriculture, The University of Jordan, Amman 11942 JORDAN

\textsuperscript{b}Food Science program, University of Missouri, Columbia, MO 65211-5160

\textsuperscript{c}Department of Clinical Nutrition and Dietetics, Faculty of Allied Health Sciences, Hashemite University, Zarqa-Jordan

\textsuperscript{d}Chemistry and Biochemistry Program, Eastern Oregon University, One University Blvd., La Grande, OR 97850-2899

\textsuperscript{e}Department of Food Science and Human Nutrition, Box 646376, Washington State University, Pullman, WA 99164-6376

*Corresponding author. Telephone: (+962 6) 5355000, Ext. 22422; Fax: (+962 6) 5355577; Email address: h.qadiri@ju.edu.jo
Visible and short wavelength near infrared (SW-NIR) diffuse reflectance spectroscopy (600-1100 nm) was evaluated as a technique for detecting and monitoring spoilage of pasteurized skim milk at 3 storage temperatures (6, 21 and 37 °C) over 3 to 30 h (control, \( t = 0 \) h) (n=3). Spectra, total aerobic plate count (APC) and pH were obtained, with a total of 60 spectra acquired per sample. Multivariate statistical procedures: principal component analysis (PCA), soft independent modeling of class analogy (SIMCA), and partial least squares (PLS) calibration models were developed for predicting the degree of milk spoilage. PCA showed apparent clustering and segregation of milk samples that were stored at different time intervals. Milk samples that were stored for 30 h or less at different temperatures were noticeably separated from control and distinctly clustered. SIMCA analysis could correctly classify 88-93% of spectra of incubated samples from control at 30 h. A PLS model with five latent variables correlating spectral features with bacterial counts and pH yielded a correlation coefficient (\( R = 0.99 \) and 0.99) and a standard error of prediction (SEP = 0.34 log_{10} CFU per mL and 0.031 pH unit) respectively. It may be feasible to use SW-NIR to detect and monitor milk spoilage rapidly and non-invasively by correlating changes in spectral features with the level of bacterial proliferation and milk spoilage.

**KEYWORDS**: SW-NIR, milk, pasteurization, spoilage, milk storage, PCA, SIMCA, PLS.
INTRODUCTION

A large number of microorganisms are identified as contaminants of milk drawn from healthy cows that cause deterioration of milk quality (Hayes et al., 2002). The udder canal and teat surface have their own microflora so that milk drawn aseptically from the cow is not sterile but contains micrococci, streptococci, and corynebacteria (Garbutt, 1997). Milk contains other bacterial contaminants from soil, water, animal feed and bedding, and animal feces. Those include Gram-negative rods (*Pseudomonas, Alcaligenes, Acinetobacter* and *Flavobacterium*), Gram-positive bacteria (*Bacillus, Clostridium, Listeria*, and *Staphylococcus*), Enterobacteriaceae; lactic acid bacteria (*Streptococcus, Lactococcus, Lactobacillus*, and *Leuconostoc*), yeasts and molds (Garbutt, 1997; Deeth et al., 2002; Giraffa, 2003).

Milk that is produced under good hygienic conditions may contain microbial numbers as low as $10^3$ colony forming unit CFU per mL, however, under poor hygienic conditions numbers may be as high as $10^5$ CFU per mL or even greater (Garbutt, 1997). Milk contains significant nutrients including amino acids, vitamins, nucleotides, inorganic salts and trace elements that make milk an ideal growth medium for a wide range of pathogenic and spoilage microorganisms. Additionally, the high water activity (0.98) and neutral pH (6.6) provide very conductive conditions for rapid growth of a wide range of microorganisms.

Temperature has an important effect on milk quality by controlling the rate of microbial spoilage. As with many other foods, milk held at higher temperature spoils rapidly. Making milk storage under refrigeration temperature (2 – 4 ºC) across the food supply chain is essential if a reasonable shelf life is to be ensured (Garbutt, 1997).

The pasteurization process kills the vegetative cells of human pathogens, viruses, and many Gram-negative microorganisms that can cause raw milk to spoil. Post-process
recontamination with psychrotrophic bacteria may occur during filling, consumer handling, or due to the improperly cleaned pasteurizers (Gruetzmacher and Bradley, 1999; Hayes et al., 2002). Psychrotrophic Gram-negative rods (particularly the genus *Pseudomonas*) are the most significant spoilage microbe in refrigerated pasteurized milk (Walker, 1988; Hayes et al., 2002). However, storage of pasteurized milk at higher temperatures is likely to result in souring via the activity of mesophilic microorganisms that survived the heat process (*Streptococcus thermophilus*) or from the activity of psychrotrophic bacteria (*e.g.* *Pseudomonas* species) that re-contaminate pasteurized milk (Garbutt, 1997; Hayes et al., 2002). These bacterial cells produce bitter flavors in the milk and cause sweet curdling, a coagulation of milk protein. Skim milk is reported to have a shorter shelf life than whole milk (Griffiths, 1989; Deeth et al., 2002) due to the relatively higher level protease activity (Deeth et al., 2002) or might also be due to the protective effect of the fat against proteolytic attack on the proteins in whole milk (Deeth et al., 2002).

Infra red (IR) spectroscopy is a nondestructive technique. Spectral features provide biochemical information regarding the molecular composition and molecular structure of cells and tissues (including bacterial cells) as well as the nature and type of molecular interactions between different cells and tissues (Choo-Smith et al., 2001). Short wavelength near infrared (SW-NIR) Spectroscopy (600-1100 nm) is widely used in food analysis, recently for quantitative detection of the microbial spoilage in chicken meat (Lin et al., 2004), for detection of spoilage in rainbow trout fillets (Lin et al., 2006) and quality management of beef and meat products (Ellis et al., 2002; Liu et al., 2003). SW-NIR spectroscopy is also an effective technique for non-invasive measurement of food constituents such as lipid (Lee et al., 1992), protein and certain carbohydrates and various quality features (such as color, cell structure, and
internal bruising) in agricultural products and food related industries. Accordingly, SW-NIR spectroscopy can be used in food quality control to monitor food safety, spoilage, and quality by acquiring metabolic snapshots and providing either a qualitative or quantitative estimate of the microbial load on a food samples.

The main objective of this study was to evaluate the feasibility of using visible and SW-NIR spectroscopy (600-1100 nm) to develop a rapid, non-invasive procedure for monitoring the spoilage of pasteurized skim milk that could be employed in industrial or retail settings.

**MATERIALS AND METHODS**

*Milk Sample Preparation*

Pasteurized skim milk was obtained from a local grocery refrigerated case and transported on ice to the laboratory the day the experiment was conducted. To examine the microbial quality of purchased skim milk; total aerobic plate count (APC) and pH measurements were carried out in triplicate. Milk was aseptically transferred to 20 mL sterile covered beakers and held at 6, 21, or 37 °C for \( t = 0, 3, 8, \) or 30 h (n=3).

*Microbial Analysis and pH Measurement*

Milk samples were tested at each storage temperature and time interval in triplicate manner for total APC and pH. For APC, ten-fold serial dilutions of the milk were prepared in 0.1% peptone water and a spread plate method was performed using tryptic soy agar (TSA) (Difco™, MD, USA). The plates were incubated for 48 h at 28 °C. The APC was enumerated as \( \log_{10} \) of the colony forming units (CFU) per mL and served as a reference value in the partial
least squares (PLS) models. pH measurement of milk samples was recorded at each time interval to 0.01 units.

**Spectra Collection**

Visible and SW-NIR spectra of milk samples were acquired using DPA-20 spectrophotometer (D-Squared Development, Inc., La Grande, OR, USA) for the control (0 h) and at each storage period. The spectrophotometer probe consists of 32 illumination fibers (600 µm in diameter) arranged in a concentric circle that 2 mm away from a central single pick-up fiber (Lin et al., 2006). Before spectral acquisition of milk samples, a dark and a reference spectrum of Spectralon (Labsphere, Inc., North Sutton, N.H., USA) was acquired. Spectralon is a thermoplastic resin with highly reflectance behavior in the NIR region (Lin et al., 2006). Spectralon was immersed to the milk samples to reflect SW-NIR light back to the pick-up fiber. Accordingly, the reference spectrum was automatically subtracted from the sample spectrum by the instrument in each measurement.

Each sample of 20 mL pasteurized skim milk was poured to pre-autoclaved 100 mL beaker. The probe was wipe and sanitized with 70% ethanol and then inserted directly into milk samples. Spectra were collected in the diffuse reflectance mode over a wavelength range from 600 to 1100 nm at 0.5 nm intervals. Triplicate samples were examined for each storage period with 20 spectra acquired at different locations within each sample (yielding a total of 60 spectra per sample). An individual spectrum was the average of 60 scans with a 200 ms exposure time for each scan. Samples spectra were obtained at the same temperature (~ 22 ºC) for all samples to control for spectral changes that could result from temperature differences during spectral collection. In addition, to reduce the effect of milk curdling on spectral patterns, each sample was aseptically mixed before spectra collection (milk curdling was expected due
Data Analysis

Delight version 3.2.1 (Textron Systems, Wilmington, MA) software was used to conduct data analysis. Data preprocessing algorithms including binning, smoothing, and second-derivative transformation were performed to reduce spectral features overlapping (Huang et al., 2002; Al-Qadiri et al., 2006). Spectral data were firstly binned by 2 cm$^{-1}$ and smoothed with a Gaussian function of 15 cm$^{-1}$ and then followed by a second derivative transformation of 15 cm$^{-1}$ gap (Lin et al., 2003; Al-Qadiri et al., 2006).

Multivariate statistical analysis procedures were performed to analyze data. Spectral data were statistically analyzed by principal component analysis (PCA), soft independent modeling of class analogy (SIMCA), and partial least squares (PLS) regression (Oust et al., 2004; Rodriguez-Saona et al., 2004; Al-Qadiri et al., 2006). PCA provides graphical representations of similarities and differences in spectral data between treatments (Martens and Naes, 1989; Goodacre et al., 1998; Kansiz et al., 1999) by removing random variation and generating natural clustering within a data set (Goodacre et al., 1998; Rodriguez-Saona et al., 2001; Nilsen et al., 2002). The first principal component (PC1) expresses the largest amount of spectral variation followed by the second PC (PC2) which explains the largest part of remaining spectral variation, and so on (Rodriguez-Saona et al., 2004). SIMCA classifies samples according to their degree of analogy to the training samples (Hampton et al., 2001-202; Al-Holy et al., 2006). PLS is widely used to establish a calibration model and to provide a correlation between reference data (actual values measured by a primary technique) with SW-NIR spectral data (predicted values) (Lin et al., 2003; Oust et al., 2004). The standard error of
prediction (SEP) was calculated to identify the predictive performance of the PLS calibration models (Lin et al., 2003):

\[
SEP = \sqrt{ \frac{\sum_{i=1}^{n} (X - Y)^2}{n-1}}
\]

where \(X\) is the measured bacterial count and \(Y\) is the predicted bacterial count and \(n\) is the number of samples.

**RESULTS AND DISCUSSION**

The mean log_{10} APC and the mean pH measurement for control milk samples \((t = 0 \text{ h})\) were 2.7 log_{10} CFU per mL and 6.66 respectively (Table 1). There was no significant increase in APC or obvious decrease in pH for milk samples stored at 6 °C even after 30 h of storage. However, results showed a significant increase in APC and a relatively distinct decrease in pH during the storage at 21 and 37 °C over 8-30 h which is caused by the growth of bacterial cells (Table 1).

Representative visible and SW-NIR spectra (600-1100 nm) of control and incubated milk samples stored at 6, 21, and 37 °C, for \(t = 3\) (A), 8 (B), and 30 h (C) are shown in Figure 1. The absorption peaks of water band are mainly due to the \(2\nu_1 + \nu_3\) overtone combination transition band of the O-H stretching and bending mode around 960 nm where \(\nu_1\) is the symmetric O-H stretch, \(\nu_3\) is the asymmetric O-H stretch, and \(\nu_2\) is the O-H binding mode \((13, 11)\).

There are apparent variations among the spectral patterns at different storage times and temperatures (Figure 1), however, it is difficult to study the variations in spectral patterns by examining raw absorption data (Al-Qadiri et al., 2006). Therefore, data pre-processing...
algorithms were employed to amplify spectral variations, with binning, smoothing and normalization conducted to decrease the overlap of spectral features and to eliminate high frequency instrumental noise and baseline shifts (Lin et al., 2003; Liu et al., 2003; Al-Qadiri et al., 2006).

Second derivative transformation was conducted to separate overlapping bands, reduce the variability between replicate spectra due to baseline offsets, and generate an estimate of the number of overlapped bands within a specific spectral region (Lee et al., 1992; Al-Qadiri et al., 2006). Spectral derivatization enhances the clarity of spectral features and improves apparent spectral resolution.

A representative second derivative transformation (15-point second order derivatization) of spectra for control ($t = 0$ h) and incubated milk samples stored at 6 (A), 21 (B), and 37 ºC (C), for $t = 3$ h, 8 h, and 30 h is shown in Figure 2. The variation among spectral patterns between the different treatments become more distinctive with spectral differences apparent in proteins region as a result of proteolytic reactions; with possible band assignments of RNH$_2$ around 1014 and 1031 nm, and NH around 1030 nm (Osborne and Fearn, 1986). Spectral variations at 744 and 960 nm can be assigned to OH of water and are also indicative of proteolysis. Other obvious spectral variations are linked to fatty acids region and result from lipolysis which occurs as milk spoils and may be associated with the band assignments of RCH=CH$_2$, C=CH around 844 nm, RCH=CH$_2$, HC=CH, C=CH around 861 nm, and CH around 901 nm (Osborne and Fearn, 1986). Both of proteolysis and lipolysis are commonly accompanied the growth of spoilage bacterial cells (Deeth et al., 2002). The spectral variations were significantly apparent in samples that were stored at 37 ºC for 30 h due to the high degree of spoilage and the rapid spoilage rate at elevated temperatures.
To differentiate between storage treatments, multivariate statistical analysis procedures (PCA, SIMCA, and PLS) were applied based upon variations in the spectral patterns. PCA is widely used to explain spectral data differences, capture related variations, and cluster samples depending upon variations in the spectral patterns that are associated with the degree of milk spoilage (Lin et al., 2003). A mean centered PCA (Figure 3) was carried out on the second derivative SW-NIR spectra over the range of 600-1100 nm. The two-dimensional PCA clustering results from SW-NIR spectra at different temperatures and times of storage demonstrated very distinct clustering and segregation between treatments. This indicates the feasibility of using visible and SW-NIR spectroscopy to discern variations in the biochemical changes and the level of deterioration resulting from proteolysis and lipolysis that result from bacterial activity that causes milk spoilage.

SIMCA analysis was performed to determine if samples from the various treatments could be classified based upon models developed for storage treatments at different temperature values and assigned to a class based upon the degree of analogy. According to SIMCA classification results (Figure 4) of control ($t = 0$ h) compared to the other storage treatments regarding storage temperature for $t = 30$ h, 55 out of 60 spectra (92%) of control were correctly classified. Table 2 shows SIMCA classification results of each storage treatment compared to the other test treatments regarding storage temperature at $t = 3$ h, 8 h, and 30 h.

PLS was used to establish a calibration model and provide a correlation between reference data (measured values) with SW-NIR spectral data (predicted values). Accordingly, a PLS prediction model was developed to quantify the bacterial counts (APC $\log_{10}$ CFU per mL) (Figure 5) and pH values (Figure 6) in milk samples at different storage periods, for example $t = 30$ h. Validation results of the PLS model with five latent variables for predicting the
bacterial counts and pH values yielded a correlation coefficient \( R = 0.99 \) and a standard error of prediction (SEP = 0.34 log\(_{10}\) CFU per mL and 0.031 pH unit) respectively indicating that an accurate quantification of microbial counts and pH values in milk samples can be determined from a PLS based prediction method.

**CONCLUSION**

The formation of metabolic byproducts, proteolysis and lipolysis caused by bacterial cell growth leads to a reduction in pH and undesirable sensory changes and these biochemical changes can be detected by visible and SW-NIR diffuse spectroscopy (600-1100 nm) to differentiate wholesome and spoiled milk samples without the necessity of enumerating bacteria. Multivariate data analytical techniques such as PCA could segregate storage treatments (6-37 °C storage for 0-30 h) with approximately 90% accuracy. Accurate quantification of microbial counts (SEP = 0.34 log\(_{10}\) CFU per mL) and pH (SEP = 0.031 pH unit) were obtained from PLS based prediction methods (five latent variables). This technique may be applicable for predicting the remaining shelf life and microbial loads in pasteurized milk. Further work is needed to investigate specific spoilage microorganisms and to precisely determine which sensory changes correlate with specific SW-NIR spectral features.

**ACKNOWLEDGMENT**

This research was supported in part by the International Marketing Program for Agriculture Commodities and Trade (IMPACT), Washington State University; Department of Nutrition and Food Technology, The University of Jordan, Amman, Jordan; USDA NRI Competitive Grants Program (NRICGP) project #2000-01617 and project #2000-01191; and Eastern Oregon University.
REFERENCES


Table 1: The mean $\log_{10}$ APC CFU per mL and the mean pH measurement for milk samples at each storage period

<table>
<thead>
<tr>
<th>Time (h)</th>
<th>Control</th>
<th>6 °C</th>
<th>21 °C</th>
<th>37 °C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 h</td>
<td>3 h</td>
<td>8 h</td>
<td>30 h</td>
</tr>
<tr>
<td>$\log_{10}$ CFU/mL*</td>
<td>2.7</td>
<td>2.9</td>
<td>3.0</td>
<td>3.7</td>
</tr>
<tr>
<td>pH**</td>
<td>6.66</td>
<td>6.64</td>
<td>6.62</td>
<td>6.58</td>
</tr>
</tbody>
</table>

*Samples were enumerated in triplicate using spread plate method on TSA and incubated for 48 h at 28 °C.

** pH was measured in triplicate.

Table 2. SIMCA classification results of each storage treatment compared to the other test treatments regarding storage temperature at $t = 3$ h, 8 h, and 30 h.

<table>
<thead>
<tr>
<th>Storage treatment</th>
<th>3 h</th>
<th></th>
<th>8 h</th>
<th></th>
<th>30 h</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>%</td>
<td>No.</td>
<td>%</td>
<td>No.</td>
</tr>
<tr>
<td>Control (0 h)</td>
<td>49</td>
<td>82</td>
<td>54</td>
<td>90</td>
<td>55</td>
</tr>
<tr>
<td>6 °C</td>
<td>50</td>
<td>83</td>
<td>53</td>
<td>88</td>
<td>56</td>
</tr>
<tr>
<td>21 °C</td>
<td>51</td>
<td>85</td>
<td>52</td>
<td>87</td>
<td>53</td>
</tr>
<tr>
<td>37 °C</td>
<td>54</td>
<td>90</td>
<td>53</td>
<td>88</td>
<td>55</td>
</tr>
</tbody>
</table>
Figure 1. Representative visible and SW-NIR spectral patterns (600-1100 nm) of control ($t = 0$ h) and milk samples stored at 6, 21, and 37 °C, for $t = 3$ h (A), 8 h (B), and 30 h (C).
Figure 2. Representative second derivative transformation of spectral patterns for control ($t = 0$ h) and milk samples stored at 6 (A), 21 (B), and 37 °C (C), for $t = 3$ h, 8 h, and 30 h.
Figure 3. Principal Components Analysis (PCA) for control ($t = 0$ h) and pasteurized milk samples stored at 6 (A), 21 (B), and 37 °C (C), for $t = 3$ h, 8 h, and 30 h.
Figure 4. SIMCA classification of control ($t = 0$ h) (A) compared to milk samples stored at 6 (B), 21 (C), and 37 °C (D), at $t = 30$ h.
Figure 5. Comparison of the measured with predicted bacterial count (log_{10} APC CFU per mL) for a PLS model (five latent variables): control ($t = 0$ h) (A) and milk samples stored at 6 (B), 21 (C), and 37 ºC (D) for $t = 30$ h.

Figure 6. Comparison of the measured and predicted pH for a PLS model (five latent variables): control ($t = 0$ h) (A) and milk samples stored at 6 (B), 21 (C), and 37 ºC (D), for $t = 30$ h.