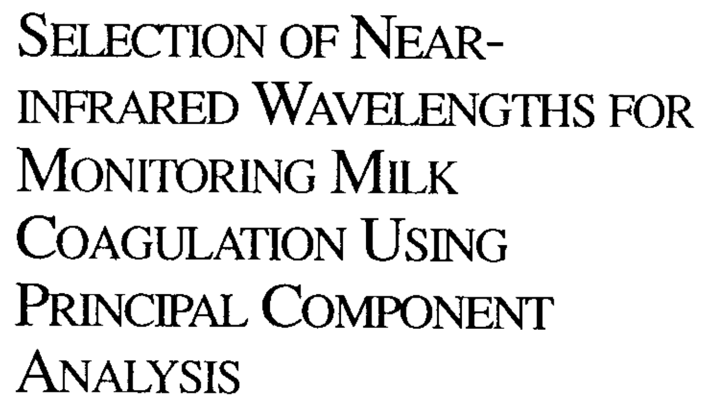

SELECTION OF NEAR-INFRARED WAVELENGTHS FOR MONITORING MILK COAGULATION USING PRINCIPAL COMPONENT ANALYSIS

D. Saputra, F. A. Payne, R. A. Lodder, S. A. Shearer

Reprinted from
Sept.-Oct. 1992 Vol. 35, No. 5 1597-1605

AMERICAN SOCIETY OF AGRICULTURAL ENGINEERS
2950 NILES RD., ST. JOSEPH, MI 49085-9659 USA

TRANSACTIONS OF THE ASAE

SELECTED CONTRIBUTIONS IN ENGINEERING FOR AGRICULTURE AND OTHER BIORESOURCE INDUSTRIES

Selection of Near-infrared Wavelengths for Monitoring Milk Coagulation Using Principal Component Analysis

D. Saputra, F. A. Payne, R. A. Lodder, S. A. Shearer
Student Member ASAE, Member ASAE, Member ASAE

Abstract

Principal component analysis was used for the selection of near infrared wavelengths for monitoring the change in macropeptide concentration of milk during the enzymatic phase of coagulation. The selection was based on the plot of the second principal component loading, the potential for monitoring phenylalanine and methionine, and the physical magnitude of the change in reflectance during coagulation. The wavelengths chosen were 1250, 1450, 1650, 1750, 1800, and 1940 nm. A multiple linear regression model was developed which correlated changes in macropeptide concentration with the first and second principal component (PC1 and PC2) scores using six selected wavelengths. This model was found to be as effective in describing the variation of the macropeptide concentration as a model that used the entire spectrum. A regression model developed using six spectral bands (± 8 nm) at the previously specified peak wavelength showed that the PC1 and PC2 components of the six spectral bands were as effective for monitoring the macropeptide concentration of coagulating milk as the entire spectrum. **Keywords.** Principal component analysis, Near infrared reflectance, Enzymes, Coagulation, Milk, Sensors.

Introduction

Milk coagulation results from the enzymatic hydrolysis of κ-casein using an enzyme extract from calf stomachs (rennet) or microbially produced enzymes and is a basic process in cheese manufacturing. Cheese yield and quality depend, among other factors, on proper control of the coagulation process.

Milk coagulation proceeds in two phases. The first phase is the hydrolysis of κ-casein into insoluble para-κ-casein and soluble macropeptide. The second phase is the subsequent aggregation of the para-κ-casein micelles into clusters. The clusters grow in size, followed by cross-linking between chains which eventually transform the milk into a gel (curd). The textural strength or firmness of the curd increases with time. These reactions depend on conditions such as pH, temperature, calcium concentration, casein concentration, enzyme activity, and temperature history of the milk.

Several instruments are available for monitoring and measuring milk coagulation, however, only a few are appropriate for in-line use in a dairy plant. van Hooydonk and van den Berg (1988) evaluated the performance of 12 dynamic-rheological-based instruments. These instruments were classified as either off-line such as the Instron®, Formagraph®, and Formascope®, or in-line such as the Gelograph®, pressure transducer sensor (PTS), Ermatic®, Vatimer®, Bendix®, and Unipan®. The disadvantage of mechanical techniques are that they perform destructive tests, are difficult to install in-line, and difficult to clean. A non-rheological method to measure the curd setting has been proposed by Hori (1985). This method is based on the changes of heat conduction from a heated wire or probe during renneting. Like the viscometric methods this method is very sensitive to the viscosity changes during aggregation. Despite the advantages and the disadvantages of the available instruments, none respond to the enzymatic phase of milk coagulation.

Clotting time is a measurement of the enzyme activity and is usually associated with the ability of the enzyme to clot milk. Clotting time is defined as the time needed from enzyme addition until a floc can be seen on a rolling bottle (Berridge, 1952). However, enzyme activity cannot be defined in terms of the rate at which milk is clotted because clotting is a two stage process (hydrolysis and aggregation), and variations in processing conditions can affect the two stages differently.

Several methods are available for the quantitative measurement of the enzymatic hydrolysis of κ-casein but none are sufficiently simple and rapid to be used for real-time analysis (Dalgleish, 1987). Available methods require lengthy preparation of a gel matrix containing urea (Chaplin and Green, 1980; and Carlson, 1982), or wet chemistry (Kjeldahl) and lengthy dialysis procedures (Wheelock and Knight, 1969; and Hyslop et al., 1979). Even high performance liquid chromatography (HPLC) takes about 40 min to produce a complete chromatogram (van Hooydonk and Olieman, 1982).

Optical changes have been reported to occur during milk coagulation and offer a promising alternative for monitoring the enzymatic reaction. Milk coagulation has been monitored through its various phases by monitoring turbidity (McMahon et al., 1984). Hardy and Fanni (1981) measured the diffuse reflectance of coagulating milk with a

Article was submitted for publication in November 1991; reviewed and approved for publication by the Food and Process Engineering Inst. of ASAE in August. Presented as ASAE Paper No. 91-6520.

This article, 91-2-211, reports the result of an investigation by the Kentucky Agricultural Experiment Station and is published with the approval of the Director. Reference to a company or trade name does not imply approval or recommendation of a product by the University of Kentucky to the exclusion of others.

The authors are **Daniel Saputra**, Graduate Research Assistant, **Fred A. Payne**, Associate Professor, Dept. of Agricultural Engineering, **Robert A. Lodder**, Assistant Professor, College of Pharmacy, and **Scott A. Shearer**, Assistant Professor, Dept. of Agricultural Engineering, University of Kentucky, Lexington.

 © 1992 American Society of Agricultural Engineers 0001-2351 / 92 / 3505-1597

color difference meter, and Korolczuk (1988) followed the refractometer signal during coagulation. Payne et al. (1990a) reported on a fiber optic sensor which monitored the diffuse reflectivity of milk during coagulation using near-infrared at a wavelength of 950 ± 5 nm.

The optical responses of milk to near-infrared reflectance (NIR) are well known and are used to measure milk constituents (Ben-Gera and Norris, 1968; Williams and Norris, 1987). Because of the promise of optical methods, it was postulated that the NIR frequency range contains responses which could be used to monitor the enzymatic phase of milk coagulation.

Researchers have developed NIR procedures to measure constituents or concentrations which would otherwise require extensive analytical procedures. Multiple linear regression (MLR) techniques have been used to analyze spectral data that usually consists of several hundred wavelengths for each sample. The spectral data at each wavelength are generally highly correlated to data at other wavelength and this correlation causes difficulties in multiple linear regression. Models developed using the multiple linear regression procedure can be easily overfitted to the data. Thus, nearly all the variance in the spectral matrix, including irrelevant information or noise, will be used in the MLR model (Beebe and Kowalski, 1987).

Principal component analysis (PCA) is a mathematical method that has been employed to analyze NIR data as a pattern recognition data reduction technique of multidimensional analysis. Beebe and Kowalski (1987) reported that PCA and partial least-squares (PLS) are preferred over multiple linear regression in analyzing spectral data. One drawback of the principal component analysis is that, even though the number of variables are reduced to three or four components, there is no corresponding reduction in the number of original variables which must be measured. A reduction in the number of original variables has an economic significance in many applications. For example, an instrument which requires six fixed wavelengths would be less costly to develop than one which required the entire near-infrared spectrum to be scanned.

Several methods have been developed for reducing the number of original variables. Jollife (1986) argued that if the number of original variables observed, p, was larger than a subset of m variables, with $m << p$, then the subset of m variables should contain virtually all the information available in all p variables. Jollife (1972) discussed several methods of reducing the number of original variables in multivariate problems, four of which utilize principal component analysis. Jollife (1973) applied five of the eight techniques to real data, including multiple linear regression analysis. He concluded that the rejection method based on principal component analysis retained too few original variables and some modification was needed in order for the method to be applicable. McCabe (1984) adopted a somewhat different approach to the selection problem. He started from the fact that principal components satisfy a number of different optimum criteria. A subset of the original variables which optimize one of these criteria is termed a set of principal variables. Jollife (1986) summarized the method of Jollife (1972, 1973) and McCabe (1984) by using an example of real data and concluded that, even though both methods show some agreement, they still retain too few original variables. For his example quoted, at least four or five of the original variables were required for appropriate prediction.

The objective of this research was to determine if principal component analysis could be used to select a minimum number of original variables (NIR wavelengths) which can be used to monitor the change in macropeptide concentration of milk during the enzymatic phase of coagulation.

Material and Methods

Milk Sample Preparation

Milk from the University of Kentucky dairy farm was pasteurized for 30 min at 63° C, separated into skimmed milk and cream, the fat content adjusted to target levels and refrigerated in a laboratory cooler at 4° C until used for testing.

Enzymatic Coagulation

The coagulating enzyme used was single strength calf rennet (Miles, Inc. Madison, WI 53701). The rennet was diluted by adding 4 mL of rennet to 36 mL of 0.1 M sodium citrate (pH 5.2) to make the final solution of 1:10 and was then refrigerated until used. Dilution with sodium citrate was performed to maintain the strength of the enzyme. The temperature of the 900 g milk sample was maintained during the test at 31 ± 0.1° C by placing the container in a water bath (Laude, Model RM20, Brinkmann Instrument, Inc., Westbury, NY 11590) as shown schematically in figure 1. The fiber optic probe and probe shield were placed in the milk and allowed to thermally equilibrate for two minute before the enzyme was added. The probe and shield were raised briefly, the enzyme was added and stirred for 5 s, and the probe and shield replaced. The two minute equilibration time was used to allow the quartz lens on the probe shield to come into thermal equilibrium with the milk.

Test Design

A test was designed to determine the effect of fat concentration and enzyme level on the selection of wavelengths. Fat levels of 1, 3, and 5% were tested at enzyme level 2 (0.20 mL rennet/kg milk). Enzyme levels 1, 2, and 3 (0.12, 0.20, and 0.28 mL rennet/kg milk, respectively) were each tested at fat level 3%. The test is shown schematically in figure 2. Each combination was replicated three times and the testing sequence was completely randomized. Skim milk was also tested at 31° C and enzyme level 2 both to quantify the concentration of macropeptide as a function of reaction time and for comparing the NIR data.

Spectral Reflectance

Spectral data were collected using a Bran+Luebbe/ Technicon Industrial Systems InfraAlyzer 500 (Technicon Corp, Tarrytown, NY) spectrometer system and a fiber-optic diffuse reflectance probe. Each spectrum was obtained from 1100 to 2500 nm in 4 nm increments giving a total of 351 original variables. Figure 3 is a diagram of the fiber-optic diffuse reflectance probe that was attached to the spectrometer to collect spectral data from milk in

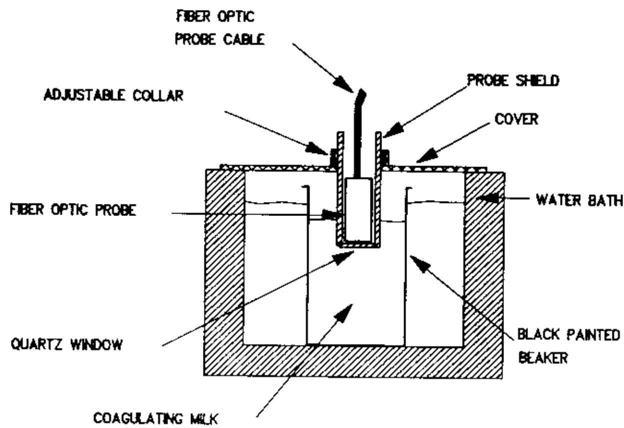

Figure 1–Schematic of experimental apparatus used to measure the diffuse reflectance of coagulating milk.

beakers. Sample beakers containing 900 g milk were placed beneath a window in a gold integrating sphere for spectral data collection. The gold sphere was inside the aluminum probe housing. Near-infrared light was directed downward into the beaker using a bundle of optical fibers at the top of the sphere, while a second optical fiber bundle on the side of the sphere directed light onto the integrating sphere wall. The reference beam directly illuminated the wall of the integrating sphere, and the sample beam directly illuminated the milk through a sapphire window. In this configuration, specular reflectance from the glass and milk surface was directed back toward the sample optical fibers while diffuse reflectance (which contained the information from light that interacted with the sample) was scattered into the integrating sphere. Spectral data were recorded as the ratio of the light intensity reflected from the milk sample to that from the reference beam.

The diffuse reflectance of the sample was measured by placing the fiber optic probe into a probe shield as shown in figure 1. The shield was fabricated of plastic pipe with a flat quartz window (3.2 mm thick) installed in the submerged end of the plastic pipe. The Technicon 500 fiber optic probe was placed in the well of the probe shield. An aluminum foil cover was used around the top of the shield to eliminate any stray light. A reflectance scan was taken prior to adding enzyme (time zero) and subsequent scans were taken at 30 s and thereafter at 3 min intervals after enzyme addition (i.e., 0, 0.5, 3, 6, 9, . . ., 57 min) for a total of 21 scans per sample. The spectrometer was connected to an IBM PS/2 MODEL 50 to record the data. The data was transferred to an IBM 3090-300E mainframe computer and analyzed using the PRINCOMP procedure of SAS (1988).

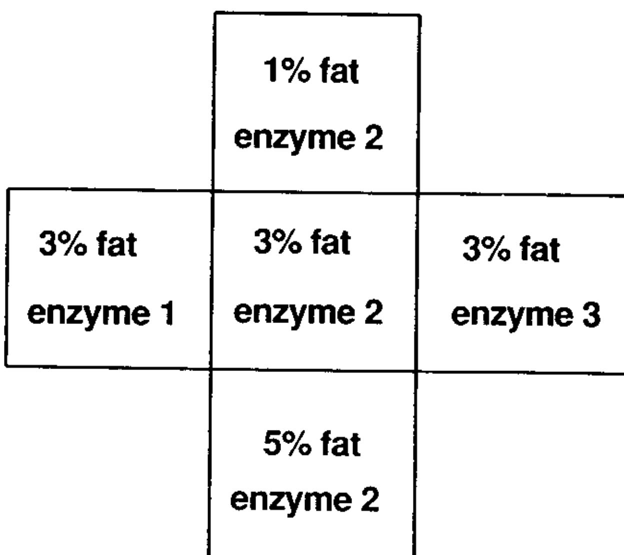

Figure 2–The experimental design layout.

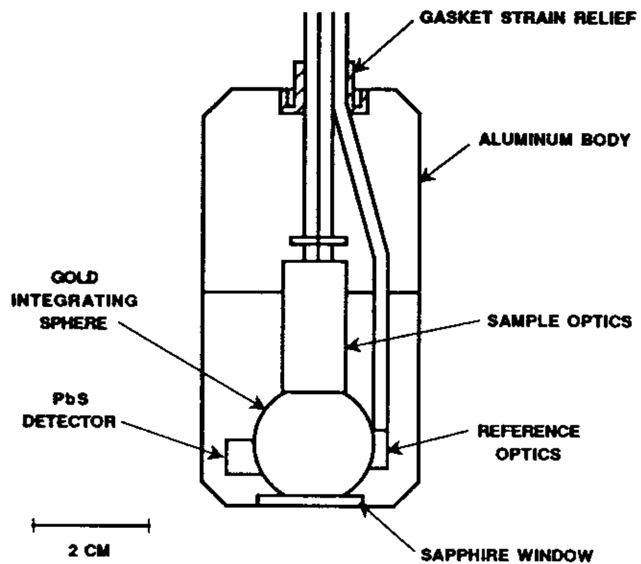

Figure 3–Schematic of the Bran+Luebbe fiber optic probe.

Macropeptide Concentration Measurement

The enzymatic hydrolysis of milk was quantified by measuring the macropeptide concentration in coagulating milk as a function of time. Skim milk was used for this procedure, since fat content was shown not to affect enzymatic reaction rates over the range of conditions tested (Payne et al., 1991). The reaction temperature was 31° C and the enzyme concentration was level 2. The macropeptide concentration, as a function of reaction time, for this sample was used to represent that for the above test at enzyme level 2 and fat concentrations of 1, 3, and 5%. Seasonal variations in milk protein content were assumed negligible for the purpose of this work.

Skim milk (800 mL) was poured into a 1000 mL flask and heated in a water bath to 31 ± 0.1° C. The milk was allowed to equilibrate for 15 min at the test temperature. The enzyme solution was added and stirred for 5 s. Nineteen samples of coagulating milk and one sample of non-enzymed milk (time-zero) were transferred into preheated 30 mL test tubes using a Repipet® dispenser (Lab Industries, Berkeley CA; accuracy 1%; reproducibility 0.1%) and were then placed back into the water bath. A test tube was randomly selected at 3-min intervals and 1 mL of 42% (w/v) trichloro acetic acid (TCA) was injected using a digital pipet (Model 4710, Eppendort®, Brinkmann Instruments, Inc., Westbury, NY, accuracy ± 0.8%) and shaken vigorously. The 2% TCA concentration was used since at this concentration all milk

proteins, except the macropeptide, are precipitated (Wheelock and Knight, 1969; Hyslop et al., 1979). The soluble macropeptide in TCA solution is referred to as non-protein nitrogen (NPN). The precipitate was filtered from the clear supernatant and stored in a freezer until analyzed.

The NPN concentration was determined by using a KJELTECR digester and distilling unit (Tecator, Sweden). NPN was calculated using AOAC procedures 2.049 and 16.036 (1975).

Data Analysis

The spectral data, which consisted of 351 reflectance observations for each scan and 21 scans per sample, were normalized to mean zero and scaled to unit variance. Normalization was utilized to give equal weighting to the observed reflectance changes throughout the spectrum (Beebe and Kowalski, 1987; Dale and Klatt, 1989). Previous experiments (Payne et al., 1990b) have shown that the normalized reflectance profiles are very similar. This technique allowed each wavelength response an equal opportunity to describe the change in chemistry independent of the magnitude of the response. The standardization was performed using the equation:

$$Z_n = \frac{X_{i,n} - \bar{X}_n}{S_n} \quad (1)$$

where

Z_n = normalized reflectance observation at wavelength n

$X_{i,n}$ = reflectance observation i at wavelength n

$\bar{X}_n$ = average reflectance observation at wavelength n

S_n = standard deviation of reflectance observations at wavelength n

The principal component analysis began with the transformation of the normalized reflectance spectra. This transformation expressed each reflectance spectrum as a point in space. The determination of the principal components was accomplished by performing an eigenanalysis on the matrix of the normalized reflectance spectra. This resulted in 351 eigenvalues and 351 eigenvectors. The relative magnitude of each eigenvalue represents the percentage of the total variance of the normalized reflectance data that occurs in the direction of the corresponding new axis or principal component. The transformation of the normalized reflectance spectra and determination of principal components were performed using the PRINCOMP procedure of SAS (1988). The Scree Test (Kim and Mueller, 1978) was used to determine the number of significant principal components by searching for a sharp drop in the amount of variation accounted for by each subsequent component. When this drop was identified, no further principal components were considered.

The coefficients of eigenvectors were scaled so that the sum of squares of the coefficients for any single component was unity. The square of these coefficients represented the relative amount of information that the corresponding reflectance observation contributed to the principal component.

Since the components have variances described by the eigenvalues matrix, the product of eigenvectors matrix to the square root of eigenvalues matrix gives the weighted relationship of the principal components to the original variables. This relationship was called the component loading. Each coefficient of the component loading is the correlation coefficient between the wavelength and the principal component (Daultrey, 1976). The plot of the loadings against the corresponding wavelengths highlight the areas of the spectrum which have the greatest influence on that component (Cowe and McNicol, 1985). Loading plots were used to determine which wavelengths had the greatest influence on the components.

The transformation of the normalized data by principal component analysis resulted in component scores. Component scores are linear combinations of the original reflectance in the direction of the principal component and can be represented by the matrix equation:

$$\{S\} = \{R\}\{E\} \quad (2)$$

where

$\{S\}$ = matrix of component scores

$\{R\}$ = matrix of original variables

$\{E\}$ = matrix of eigenvectors

Component scores of matrix $\{S\}$ as a regressor were regressed against NPN using multiple linear regression (MLR) to develop a Prediction equation.

Basically, in principal component analysis the reflectance spectra were transformed or concentrated into their most dominant dimensions and used as a regression factor instead of the whole spectra. Concentrating the data to their most dominant dimensions also produces an estimated, near infrared, loading spectra which gives various diagnostic checks that reveal abnormalities in the data, and estimates range of variations and level of measurement noise (Martens and Naes, 1987).

Results and Discussion

Typical reflectance spectra for skim milk, 1% fat, and 5% fat using enzyme level 2 are shown in figures 4, 5, and 6, respectively. These figures show that the reflectance spectra for wavelengths greater than 2100 nm were sensitive to fat concentration. Figure 4 shows that the reflectance spectrum for skim milk after 2100 nm changed very little with time but the spectra for the higher fat contents (figs. 4 and 5) show a significant change with time. For these reasons the principal component analysis was focused on the spectrum from 1100 to 2100 nm.

Principal component analysis was conducted on each of the five treatments shown in figure 2. Each treatment consisted of three replicated samples and 21 spectral scans for each sample for a total of 63 spectral scans. PC1 accounted for more than 90% of the variation and PC2 accounted for the range of 1.2% to 2.95% for all five treatments as listed in Table 1. The Scree test shown that only PC1 and PC2 significant to the test.

Figures 7 and 8 show the typical plot of PC1 loading versus wavelength for 1% and 5% fat content at enzyme

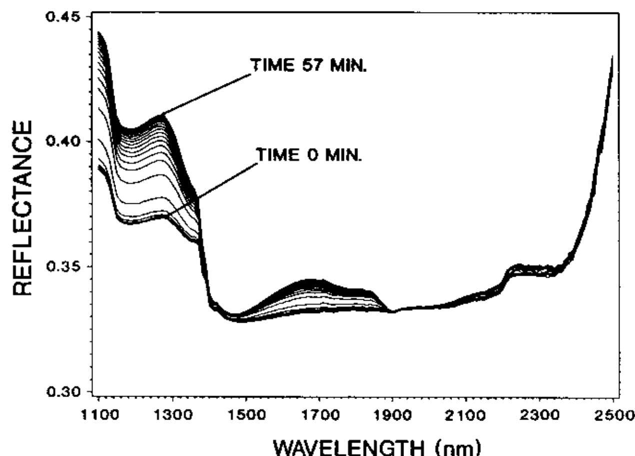

Figure 4–Reflectance spectra for coagulation of skim milk using enzyme level 2 at 31° C.

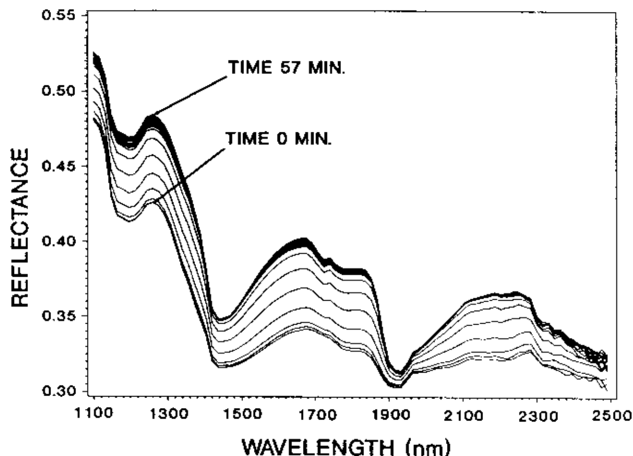

Figure 6–Reflectance spectra for coagulating milk containing 5% fat and using enzyme level 2 at 31° C.

level 2. The PC1 loading was almost flat across the spectrum except around 1940 nm (water) and after 2300 nm. The lack of any localized loading caused by individual reflectance bands, apart from the effect of moisture, suggest that this component measures an overall change. Cowe and McNicol (1985) stated that a principal component in which the loadings are approximately equal can be considered as a mean of all the reflectance values. An explanation for the high proportion of the total spectral variation to PC1 is that the rate of change of reflectance with time at all wavelengths was about the same. It is postulated that this resulted because as coagulation proceeds the mean light pathlength was reduced uniformly for all wavelengths by the increasing turbidity associated with coagulation.

PC2 was thought to show effects independent of the change in turbidity of the coagulating milk. The concentration of macropeptide increases as hydrolysis proceeds (Carlson, 1982; Dalgheis, 1987; Green, 1972). The increasing macropeptide concentration in the liquid surrounding the casein particles may result in absorbance changes of near-infrared light for selected wavelengths.

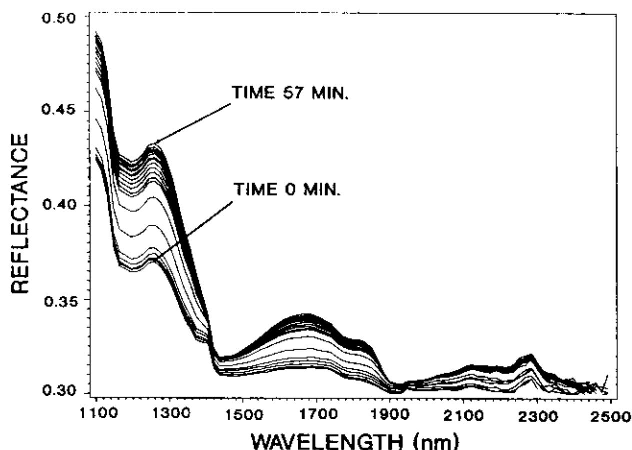

Figure 5–Reflectance spectra for coagulating milk containing 1% fat and using enzyme level 2 at 31° C.

Figures 9 and 10 show that there are two points of contrast for the PC2 loading. The positive loading around wavelengths 1450 nm and 1940 nm were considered to result from water (Cowe and McNicol, 1985; Williams and Norris, 1987). The negative loadings at 1100 to 1300 nm and 1600 to 1850 nm may be related to macropeptide concentration. The milk clotting enzyme splits up κ-casein at the junction of para-κ-casein and macropeptide at the bond between phenylalanine residue 105 and the methionine residue 106 (Dalgheis, 1987; Carlson, 1982). The absorption spectra of phenylalanine, methionine and their second derivative (Williams and Norris, 1987) show second derivative peaks for wavelengths between 1150 and 1300 nm, between 1600 and 1850 nm, and after 2100 nm similar to the PC2 loadings. The concentration of NPN typically increases during enzymatic coagulation from 0.4% to 0.5%. NPN concentration was thus considered a potential factor in changes in the absorbance of light.

The positive loadings after 2100 nm were not considered in the wavelength selection for two reasons. First, as can be seen from figures 7, 8, 9, and 10, the loadings at these wavelengths were mainly contributed by noise, and second, the reflectance of the wavelengths after 2100 nm were sensitive to fat concentration as mentioned earlier. Figure 4 shows that the reflectance spectrum for skim milk after 2100 nm changed very little with time but for the higher fat content (figs. 5 and 6) the reflectance spectrum after 2100 nm shows a significant change with time.

TABLE 1. Percent of data variation explained by the principal component analysis

	Variation Explained		
Treatment	PC1 (%)	PC2 (%)	Total (%)
1% Fat, enzyme 2	90.34	2.95	93.29
3% Fat, enzyme 2	95.33	2.35	97.68
5% Fat, enzyme 2	95.74	1.23	96.97
3% Fat, enzyme 1	96.07	1.59	97.66
3% Fat, enzyme 3	93.90	2.13	96.03

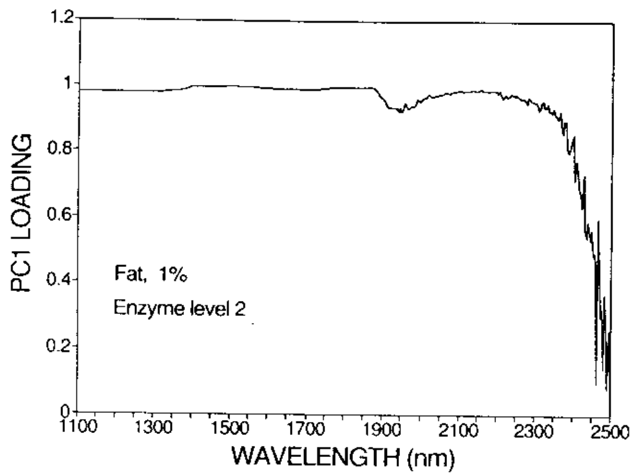

Figure 7–Plot of PC1 loading for coagulating milk (1% fat; enzyme level 2).

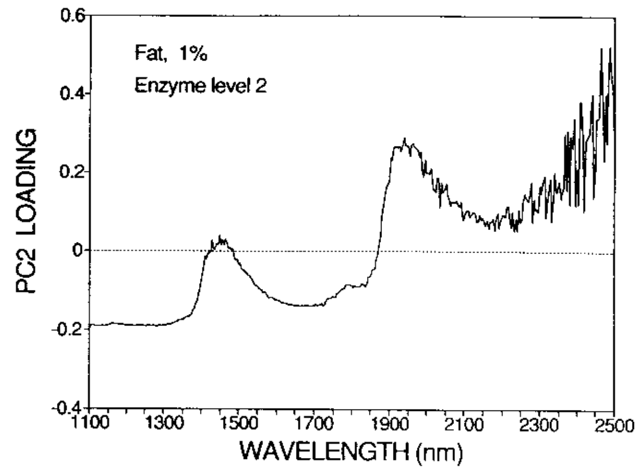

Figure 9–Plot of PC2 loading for coagulating milk (1% fat; enzyme level 2).

The macropeptide concentration (NPN) data were utilized using the non-linear regression procedure of SAS (1988). The NPN values were normalized with NPN equal to zero at time zero and equal to one at time 57 min. A NPN prediction equation was developed by fitting the first-order kinetic model (Carlson, 1982; van Hooydonk and Walstra, 1987) to the data to give the following equation:

$$NPN_e = \left(1 - e^{-t/\tau}\right) \quad (3)$$

where

τ = macropeptide generation time constant (min)

NPN_e = predicted normalized macropeptide concentration

t = enzymatic reaction time (min)

The macropeptide generation time constant at the conditions tested was 8.3 min. The mean initial NPN concentration was 0.45% and the mean final concentration (t = 57 min) was 0.54%. Figure 11 shows the plot of the hydrolysis data and NPN_e as a function of coagulation time. The scatter of the hydrolysis data resulted from inherent measurement errors associated with the Kjeldahl procedure and equipment and the relatively small change in magnitude of NPN.

Equation 3 was used to generate NPN_e data for regression against principal component scores rather than using the raw data means for NPN which included scatter. This was done to provide an equal number of data point between the NPN_e and the principal component scores for the skim milk. Additionally, equation 3 was used to generate the NPN_e data for the 1, 3, and 5% fat test at enzyme level 2. The regression of PC1 and PC2 scores against NPN_e is summarized in Table 2 for the three treatments using enzyme level 2. Model standard errors of 0.09 (normalized NPN_e unit, i.e., 9% of the difference in NPN between time zero and time 57 min) resulted on all tests except the 5% fat treatment which gave an R^2 of 0.83 and model standard error of 0.1. Figure 12 shows the plot of predicted NPN against NPNe using the regression model and coefficients generated for the 5% fat content and enzyme level 2. This shows that principal component

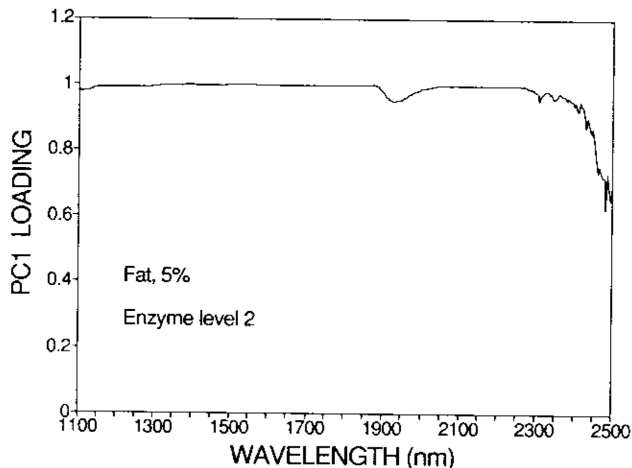

Figure 8–Plot of PC1 loading for coagulating milk (5% fat; enzyme level 2).

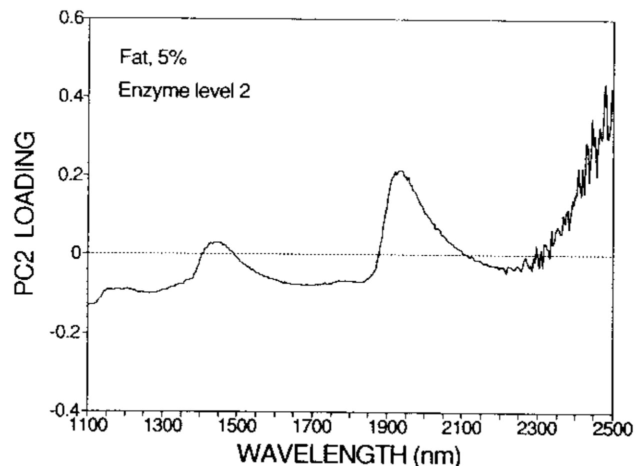

Figure 10–Plot of PC2 loading for coagulating milk (5% fat; enzyme level 2).

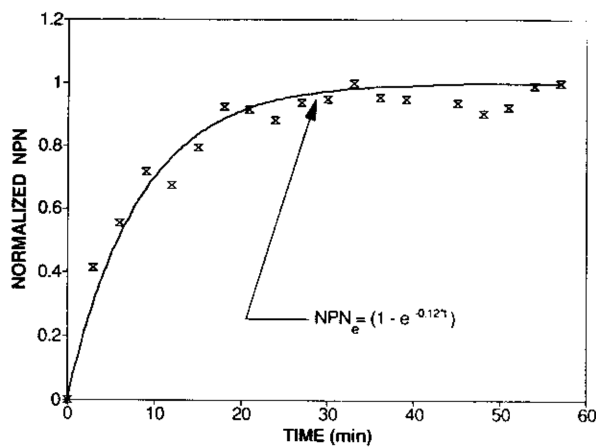

Figure 11–A plot of the normalized means for NPN_e data as a function of coagulation time.

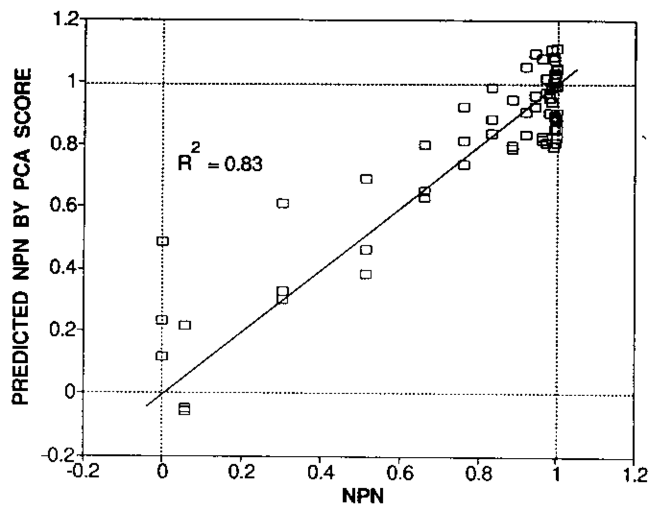

Figure 12–Plot of NPN predicted by a MLR model using principal component scores for all wavelengths as function of NPN_e for coagulating milk containing 5% fat and enzyme level 2 at 31° C (data points consist of three replicates and 21 scans/replicate).

analysis of NIR spectrum offers a potential for predicting the κ-casein hydrolysis. The scatter at time zero was considered the result from both the dilution effect of adding the aqueous enzyme solution to the milk and possibly a cooling effect of the quartz lens on the probe when removed to add the enzyme.

Wavelength Selection for Monitoring Milk Coagulation

The 1450 nm and 1940 nm wavelengths were selected because of the positive loading and their relation to water absorption. Beside the water effects, two wavelength ranges from 1100 to 1300 nm and 1600 to 1850 nm were identified as having potential for monitoring the change in macropeptide during the milk coagulation. An individual wavelength of 1250 nm from the range of 1100 to 1300 nm and wavelengths of 1650, 1750, and 1800 nm from the 1600 to 1850 nm range were selected because of the absolute value of PC2 loading, their potential for monitoring phenylalanine and methionine, and the physical magnitude of the change in reflectance during coagulation.

Another principal component analysis was performed on the spectral data of these six wavelengths. Table 3 shows that the first two principal components for the six wavelengths chosen explained >99% of the variation. The component scores for PC1 and PC2 were then regressed against NPN_e to develop a prediction model. The regression model was shown to be as effective in predicting NPN as the above regression model based on the entire spectra. Figure 13 shows the plot of predicted NPN as a function of NPN_e for the MLR model for the 5% fat content, enzyme level 2 treatment.

Narrow spectral bands filters are commonly applied to dedicated NIR instruments. To determine if spectral bands could be used to predict NPN, the principal component analysis was also conducted on six 16 nm wide spectral band. The reflectance data for the six bands were generated by averaging the reflectance observations collected within ± 8 nm of the center wavelengths selected previously. Table 2 shows that principal component scores for the six spectral bands were a good regressor to NPN_e. The regression coefficients, R^2 and C.V. for the six bands were

TABLE 2. Summary of the multiple linear regression of PC1 and PC2 against NPN_e for the three fat level treatments using enzyme level 2

Treatment	Wavelengths	β_0	$S_{\beta 0}$	β_1	$S_{\beta 1}$	β_2	$S_{\beta 2}$	R^2	C.V. (%)
1% Fat	All	0.798*	0.014	0.017*	0.001	0.011†	0.004	91.6	11.6
	Six‡	0.798*	0.015	0.122*	0.006	0.080†	0.033	90.6	12.2
	Six bands§	0.798*	0.015	0.122*	0.006	0.087†	0.035	90.5	12.3
3% Fat	All	0.798*	0.011	0.016*	0.001	−0.017*	0.004	91.9	11.2
	Six‡	0.798*	0.012	0.121*	0.005	−0.104*	0.031	91.3	11.6
	Six bands	0.798*	0.012	0.121*	0.005	−0.109*	0.031	91.4	11.5
5% Fat	All	0.798*	0.016	0.015*	0.001	0.041*	0.008	83.4	16.1
	Six‡	0.798*	0.012	0.112*	0.005	0.325*	0.035	90.4	12.2
	Six bands§	0.798*	0.012	0.112*	0.005	0.335*	0.035	90.8	12.0

* Significant at α level of 1%.

† Significant at α level of 5%.

‡ Wavelength of 1250, 1450, 1650, 1750, 1800, and 1940 nm.

§ The spectral bands of ±8 nm with mid-wavelength of 1250, 1450, 1650, 1750, 1800, and 1940 nm.

TABLE 3. Proportion of total Principal component (PC1 + PC2) to the variation of hydrolysis of κ-casein using all wavelengths, six wavelengths, and six spectral bands

	PC1 + PC2 (%)		
Treatment	All Wavelengths	Six Wavelengths*	Six Spectral Bands†
1% Fat Enzyme 2	93.29	99.81	99.87
3% Fat Enzyme 2	97.68	99.89	99.89
5% Fat Enzyme 2	96.97	99.71	99.71
3% Fat Enzyme 1	97.66	99.84	99.85
3% Fat Enzyme 3	96.03	99.53	99.56

* 1250, 1450, 1650, 1750, 1800, and 1940nm.

† The spectral band of ± 8 nm with mid-wavelength of 1250, 1450, 1650, 1750, 1800, and 1940 nm.

nearly identical to those for the six selected wavelengths. Thus principal component analysis shows that six narrow spectral bands can be as effective for monitoring the enzymatic phase of milk coagulation as the entire spectrum. Future work will be directed toward a more detailed investigation of the relationship of spectral bands with enzymatic hydrolysis.

Conclusion

- The near infrared wavelength ranges of 1200 to 1300 and 1600 to 1800 nm and wavelengths of 1940 and 1450 nm show potential for monitoring the enzymatic coagulation of milk.

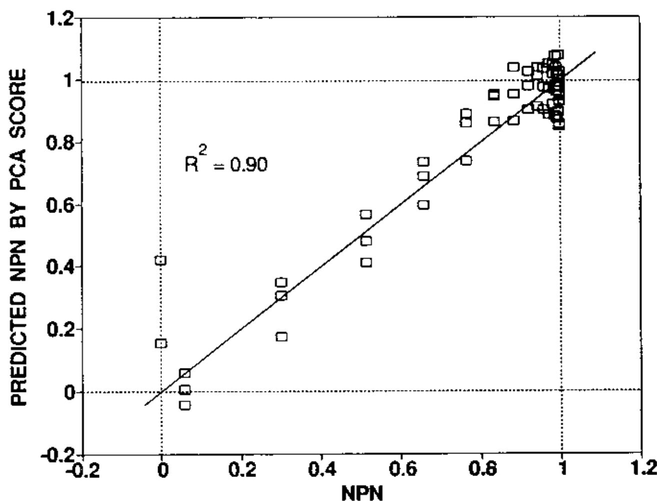

Figure 13–Plot of NPN predicted by a MLR model using principal component scores for six selected wavelengths as function of NPN_c for coagulating milk containing 5% fat and enzyme level 2 at 31° C (data points consist of 3 replicates and 21 scans/replicate).

- A prediction model based on the regression of principal component scores for six wavelengths (1250, 1450, 1650, 1750, 1800, and 1940 nm) against macropeptide concentration was found to be as effective in predicting the enzymatic hydrolysis of milk as a model based on the entire near infrared reflectance spectrum.
- A prediction model based on the regression of principal component scores for six 16 nm width spectral bands against macropeptide concentration was found to be as effective in predicting the enzymatic hydrolysis of milk as a model based on the entire near infrared reflectance spectra.

References

A.O.A.C. 1975. *Official Methods of Analysis*, 12th Ed. P.O. Box 540, Benjamin Franklin Station, Washington, DC 20044.

Beebe, K. R. and B. R. Kowalski. 1987. An introduction to multivariate calibration and analysis. *Analytical Chemistry* 59(17):1007A-1017A.

Ben-Gera, I. and K. H. Norris. 1968. Influence of fat concentration on the absorption spectrum of milk in the near infrared region. *Israel J. Agric. Res.* 18(3):117-132.

Berridge, N. J. 1952. An improved method of observing the clotting of milk containing rennin. *J. of Dairy Res.* 18:328-329.

Carlson, A. 1982. The enzymatic coagulation of milk. Ph.D diss., University of Wisconsin, Madison.

Chaplin, B. and M. L. Green. 1980. Determination of the proportion of κ-casein hydrolyzed by rennet on coagulation of skim milk. *J. of Dairy Res.* 47:351 - 353.

Cowe, I. A. and J. W. McNicol. 1985. The use of principal components in the analysis of near-infrared spectra. *Applied Spectroscopy* 39(2):257-265.

Dale, J. M. and L. N. Klatt. 1989. Principal component analysis of diffuse near-infrared reflectance data from paper currency. *Applied Spectroscopy* 43(8):1399-1405.

Dalgleish, D. G. 1987. The enzymatic coagulation of milk. In *Cheese: Chemistry, Physic, and Microbiology*, ed. P. F. Fox, Vol. 1:63-96. London: Elsevier Applied Science.

Daultrey, S. 1976. Principal Components Analysis. CATMOG 8 (Concept and Techniques in Modern Geography No. 8). Norwich: Geo Abstract Ltd.

Green, M. L. 1972. On the mechanism of milk clotting by rennin. *J. Dairy Res.* 39:55-63.

Hardy, J. and J. Fanni. 1981. Application of reflection photometry to the measurement of milk coagulation. *J. of Food Sci.* 46:1956-1957.

Hori, T. 1985. Objective measurement of the process of curd formation during rennet treatment of milks by the hot wire method. *J. of Food Sci.* 50:911-917.

Hyslop, D. G., T. Richardson, and D. S. Ryan. 1979. Kinetics of pepsin-initiated coagulation of κ-casein. *Biochimica Biophysica Acta* 566(2):390-396.

Jollife, I. T. 1986. *Principal Component Analysis*. Springer Series in Statistics. New York: Springer-Verlag, Inc.

————. 1973. Discarding variables in a principal component analysis. 11: Real data. *Applied Statistics* 22:21-31.

————. 1972. Discarding variables in a principal component analysis. 1: Artificial data. *Applied Statistics* 21:160-173.

Kim, J and C. W. Mueller. 1978. Factor Analysis. Statistical Methods and Practical Issues. SAGE University Paper. Beverly Hills, CA.

Martens, H. and T. Naes. 1987. Multivariate calibration by data compression. In *Near Infrared Technology in the Agriculture and Food Industries*, eds. P. Williams and K. Norris. St. Paul, MN: American Association of Cereal Chemist, Inc.

Korolczuk, J. 1988. Refractometric signal evolution during rennet coagulation of milk. *Neth. Milk and Dairy* J. 42:313-322.

McCabe, G. P. 1984. Principal variables. *Technometrics* 26(2):137-144.

McMahon, D. J., R. J. Brown and C. A. Ernstrom. 1984. Enzymic coagulation of milk casein micelles. *J. Dairy Sci.* 67:745-748.

Payne, F. A., C. L. Hicks and Pao-Sheng Shen. 1993. Predicting optimal cut-time of coagulation milk using diffuse reflectance. *J. Dairy Sci.* (76, In press).

Payne, F. A., S. Madangopal, C. L. Hicks and S. Shearer. 1990a. Fiber optic milk coagulation sensor for cut-time detection. *Food Processing Automation. Proceedings of the 1990 Conference.* St. Joseph, MI: ASAE.

Payne, F. A., D. Saputra, S. A. Shearer and R. A. Lodder. 1990b. Near infrared reflectance of milk during enzymatic coagulation. ASAE Paper No. 90-6075. St. Joseph, MI: ASAE.

SAS . 1988. SAS Institute Inc., Release 5.18, Cary, NC.

van Hooydonk, A. C. M. and G. van den Berg. 1988. Control and determination of the curd-setting during cheese making. Bulletin of the International Dairy Federation No. 225:210.

van Hooydonk, A. C. M. and P. Walstra. 1987. Interpretation of the kinetics of the renneting reaction in milk. *Neth. Milk and Dairy J.* 41:19-47.

van Hooydonk, A. C. M. and C. Olieman. 1982. A rapid and sensitive high performance liquid chromatography method following the action of chymosin in milk. *Neth. Milk and Dairy J.* 36:153.

Wheelock, J. V. and D. J. Knight. 1969. The action of rennet on whole milk. *J. of Dairy Res.* 36:183-190.

Williams, P. and K. Norris. 1987. *Near Infrared Technology in the Agriculture and Food Industries.* St. Paul, MN: American Association of Cereal Chemist, Inc.