Linearization and Scatter-Correction for Near-Infrared Reflectance Spectra of Meat

P. GELADI, D. MacDOUGALL, and H. MARTENS*

Umeå University, Umeå, Sweden (P.G.); Meat Research Institute, Bristol, UK (D.M.); and Norwegian Food Research Institute, Ås, Norway (H.M.)

This paper is concerned with the quantitative analysis of multicomponent mixtures by diffuse reflectance spectroscopy. Near-infrared reflectance (NIRR) measurements are related to chemical composition but in a nonlinear way, and light scatter distorts the data. Various response linearizations of reflectance (R) are compared (R with Saunderson correction for internal reflectance, log $1/R$, and Kubelka-Munk transformations and its inverse). A multi-wavelength concept for optical correction (Multiplicative Scatter Correction, MSC) is proposed for separating the chemical light absorption from the physical light scatter. Partial Least Squares (PLS) regression is used as the multivariate linear calibration method for predicting fat in meat from linearized and scatter-corrected NIRR data over a broad concentration range.

All the response linearization methods improved fat prediction when used with the MSC; corrected log $1/R$ and inverse Kubelka-Munk transformations yielded the best results. The MSC provided simpler calibration models with good correspondence to the expected physical model of meat. The scatter coefficients obtained from the MSC correlated with fat content, indicating that fat affects the NIRR of meat with an additive absorption component and a multiplicative scatter component.

Index Headings: **Light scatter; Nonlinear; Multivariate; Calibration; Near-Infrared; Reflectance.**

INTRODUCTION

Two important sources of inaccuracy in the determination of chemical composition by diffuse reflectance spectroscopy are light scatter and nonlinearity of instrument response.

Rapid measurement of the chemical composition of foods and food products by diffuse reflectance spectroscopy in the near-infrared (NIR) region (0.8–2.6 μm) has been in use for some time, and its application to meat is receiving increasing attention.[1-3] Its basis is that the major food constituents—water, fat, protein and carbohydrates—have distinct, although overlapping, absorption bands in the NIR region of the spectrum. Because most foods are nontransparent and many are highly scattering, the mode of measurements is usually reflectance (NIRR), although diffuse transmission (NIRT) is also possible.

Raw meat can be regarded chemically as a linear combination of two components, muscle tissue and fat tissue, each with relatively constant composition of fat, water, and protein. Since the two components add up to 100%, meat samples chemically vary mainly along a single linear dimension, and ideally should be suitable for univariate calibration, i.e., one using only one wavelength. In practice, the conventional NIRR unit requires a more complex calibration using many different wavelengths.

Received 29 June 1984.

* Author to whom correspondence should be addressed.

Since meat and other foods are heterogenous, they must be homogenized prior to measurement. This procedure greatly affects reflectance because the intensity of scattered light depends on, among other things, particle size distribution.

In severe cases as much as 99% of the variance in NIRR spectra may be caused by systematic scatter noise.[4] Nevertheless, commercial NIR instruments work surprisingly well using linear calibration based on data from several wavelengths. Scatter noise affects the apparent absorbance in a nonlinear way, but the linear *additive* calibration methods employed have the ability to model nonlinear relationships by increasing the multivariate complexity as long as the noise effects are not strictly *multiplicative*. However, the higher complexity increases the number of calibration coefficients that have to be estimated, which in turn increases the requirement for independent calibration samples. If scatter and other systematic noise types such as specular reflectance, could be corrected for mathematically, one might expect simpler, cheaper, and more accurate calibration of NIR instruments.

In NIRR a ratioing technique is available for empirically reducing light scatter effects from scanning NIR instruments.[5] Norris[6] gives a physical explanation of the method. Usually it implies a division of the second derivative of the NIR spectrum at a wavelength showing exceptionally high absorption by the constituents in question, by the second derivative at a wavelength where all the major constituents show about equal absorptions.

The aim of the present study is to describe a new concept, Multiplicative Scatter Correction (MSC), for treating light scatter variations and to evaluate it for meat NIRR data. Application of this concept gave significant improvements for wheat,[4] and preliminary results indicated that it had potential for meat.[7] The MSC concept is an extension of empirical ratioing,[5] using many wavelengths and allowing both additive and multiplicative optical effects. It is tested presently for reflectance and several transforms, including apparent absorbance and that of Kubelka and Munk, procedures intended for linearizing instrument response.

Notation. The standard spectroscopic notation of single capital letters is used to represent the spectral data in different units (R, A, (K/S), etc.). The MSC transformation is expressed in general form, with letter X representing NIRR data in any of these units.

Data in these units after MSC transformation are represented by letters with a subscript 1 (R_1, A_1, $(K/S)_1$, etc.).

To give a simple explanation of multivariate calibration, prediction, and transformations of Beer's law, one

0003-7028/85/3903-0491$2.00/0

TABLE I. Wavelengths used in the Technicon InfraAlyzer 400.

Wavelength, nm	Typical for
1445	Water
1680	
1722	
1734	Protein
1759	Fat
1778	
1818	
1940	Water
1982	
2100	Starch
2139	
2180	Protein
2190	
2208	
2230	
2270	
2310	Fat
2336	
2348	

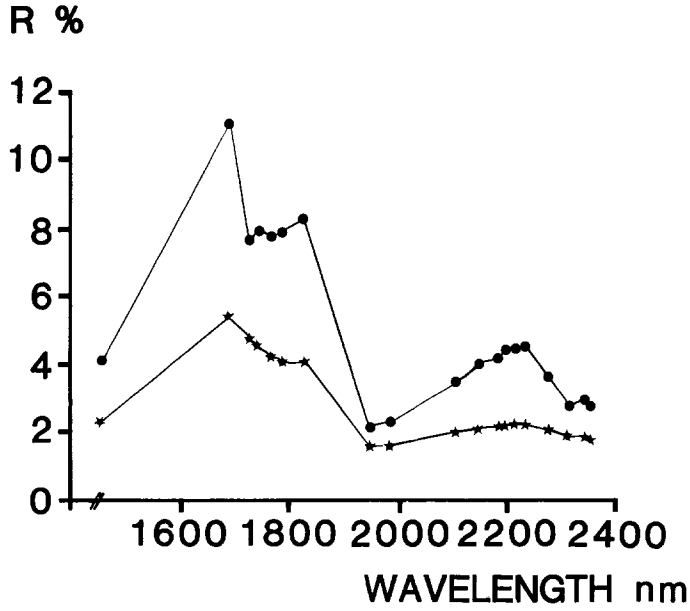

FIG. 1. NIRR spectra of a high-fat (●) and a low-fat (*) sample at the 19 wavelengths of the InfraAlyzer 400 spectrometer.

uses a more statistically oriented general notation where x_{ik} represents spectral data in any unit (R, R_1, A, A_1, etc.) for sample i at wavelength k.

Chemical concentrations are in spectroscopic formulae denoted by c (when only fat is considered) or by c_j (when two or more constituents ($j = 1, 2, \ldots$) are considered). In the statistical formulae the fat concentration in sample i is denoted by c_i; $\hat{c}_i$ (c_i-"hat") represents the fat percentage predicted mathematically from the NIRR data.

EXPERIMENTAL—SAMPLES AND INSTRUMENTATION

Raw bovine and porcine meat samples were ground, mixed, and then homogenized in an Ultra-Turax grinder for various lengths of time. NIRR spectra of the homogenized samples in open sample cups with their surface smoothed to different degrees with a spatula were measured on a Technicon InfraAlyzer 400 at 19 distinct wavelengths in the region 1400 to 2400 nm (Table I). The InfraAlyzer 400 transforms reflectance to apparent absorbance and stores the data on magnetic tape. Typical spectra of low- and high-fat samples are shown in Fig. 1.

The samples were analyzed by standard wet chemical techniques for fat percentage by extraction with an organic solvent (Foslet). In this paper fat is the only chemical constituent discussed.

Data Set I. The first data set was constructed to provide a study of the influence of homogenization and surface smoothness on the reflectance spectra of high-fat and low-fat samples.

Data Set II. The second data set consisted of binary

TABLE II. Number of meat samples at different treatments.

Set 1: Effect of homogenization and surface quality for lean and fat meat (total no. of samples: 24)

		Homogenization		
	Surface	Unhomog.	1× Homog.	2× Homog.
Pork (3% fat)	Rough	2	2	2
	Smooth	2	2	2
Beef (17% fat)	Rough	2	2	2
	Smooth	2	2	2

Set 2: Effect of changed chemical composition at two different levels of homogenization (total no. of samples: 48)

Sample ratio	100/0	85/15	75/25	65/35	50/50	35/65	25/75	15/85	0/100
% fat	26								3
Homogenization									
1 ×	4	2	4	2	4	2	2	2	2
2 ×	4	2	4	2	4	2	2	2	2

Set 3: Commercial meat cuts (total no. of samples: 60)

	Beef			Pork	
Grading	I	II	III	I	II
% fat	0–7	7–17	17–30	0–7	7–30
No.	12	12	12	12	12

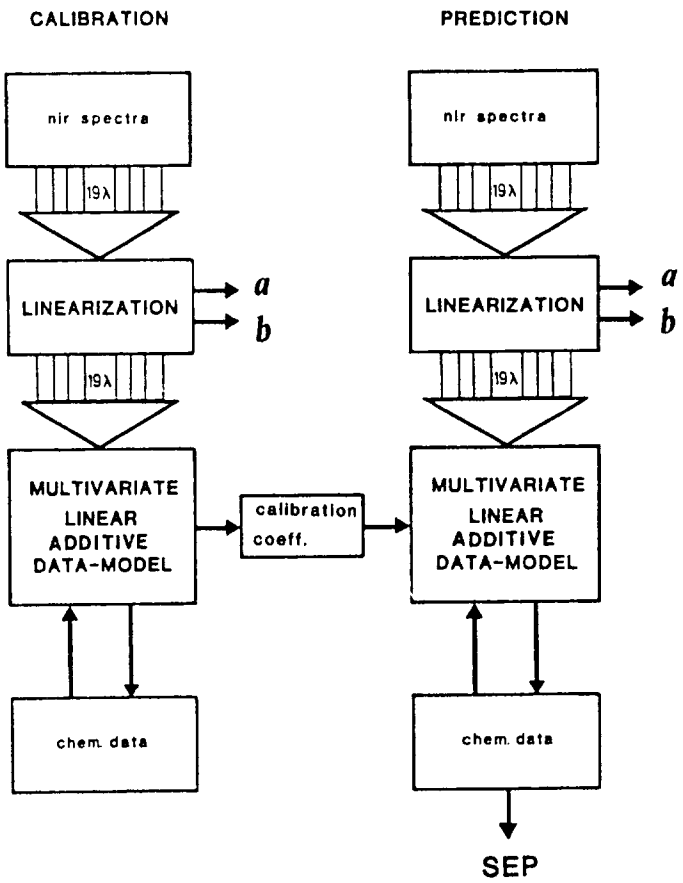

FIG. 2. Outline of the steps in multivariate calibration and prediction. Based on the input of NIR reflectance and chemical data in a calibration set of samples the calibration constants are estimated statistically. These are used in predicting chemical data in new samples. The spectrum of each sample is subjected to some linearization which may produce constants a and b.

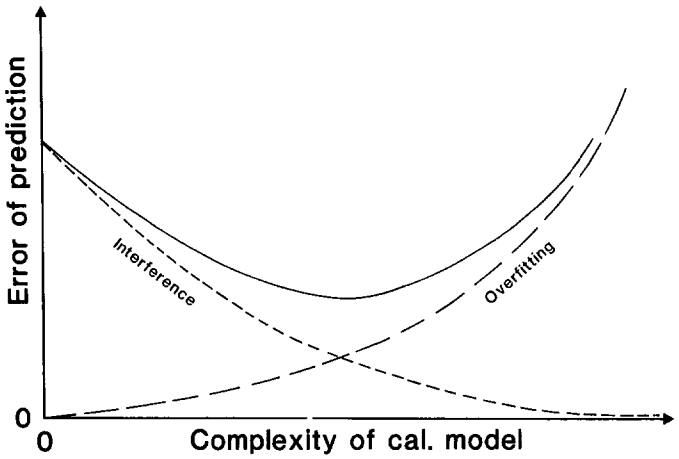

FIG. 3. Theoretical illustration of model optimization. Error of prediction as function of complexity of the calibration model (its dimensionality). Dotted curve: Systematic error due to interference from the complex chemical composition and optical properties of the samples. Broken curve: Random error from statistical calibration on a finite sample set, leading to over-fitting. Solid curve: Resulting prediction error.

mixtures of high-fat and low-fat meat constructed for study of the linearity of the spectroscopic response to changing fat content.

Data Set III. The third set was selected from an industrial production line and represents what might be expected in normal analysis situations.[7] Details of the sets are shown in Table II.

Computer software for data storage, transformations, calibration and prediction was developed by the authors on a Hewlett-Packard 85 microcomputer in BASIC language.

THEORY

Multivariate Calibration and Prediction. In NIRR spectroscopy, concentration c_i of a chemical constituent (e.g., fat) in sample no. i is modelled as a multivariate function of the spectral measurements x_{ik} at the different wavelengths $k = 1, 2, \ldots, K$. For the present type of NIRR instruments a linear additive calibration model is used:

$$c_i = g_0 + \sum_{k=1}^{K} g_k \cdot x_{ik} + e_i \qquad (1)$$

where g_0 and g_k represent calibration coefficients, e_i represents the statistical residuals caused by measurement noise in c_i and x_{ik} and by deviations from this linear model.

The calculation consists of the following steps, as shown in Fig. 2:

Calibration. Calibration coefficients g_0 and g_k are estimated statistically by Indirect Calibration, relating the chemical variable c_i to the different NIRR wavelengths x_{ik}, $k = 1, 2, \ldots, K$ in a set of calibration samples $i = 1, 2, \ldots, I$ by a multivariate linear additive regression method. One uses several wavelengths in the final calibration results in order to compensate for unknown, but systematic interferences in the spectral data (interferences due to other chemical constituents with overlapping spectra, interferences due to light scatter effects, and interferences due to, e.g., temperature variations between samples). In this way it is possible to obtain quantitative results for c_i from "fuzzy," nonselective data x_{ik}, such as NIRR data. The Partial Least Squares (PLS) regression method was used here because of its good statistical properties.[8,9] It is similar to the Stepwise Multiple Linear Regression (SMLR) conventionally used in this type of NIRR instrument, but uses *all* relevant wavelengths and provides better error warnings. The algorithm is explained in detail by Martens *et al.*[10] While SMLR uses one extra *wavelength* to model each linear interference phenomenon, PLS regression instead computes an extra "factor" (a new linear *combination* of *all* the 19 measured wavelengths) for each such phenomenon.

In SMLR the coefficients g_0 and g_k, $k = 1, 2, \ldots, K$ are estimated by ones fitting the fat data to the K-selected wavelengths using Eq. 2 directly to model c_i from given x_{ik}. In PLS the coefficients g_0 and g_k, $k = 1, 2, \ldots, K$ are obtained by a mathematical transformation of the PLS factor solution[10] that models both c_i and x_{ik}. The statistical details of the linear multivariate modelling statistics are not the subject of the present paper, which instead focuses on how to simplify and improve

the linear modelling by mathematical linearity transformations of the NIRR data.

Prediction. The obtained linear calibration coefficients $\hat{g}_0$ and $\hat{g}_k$, $k = 1, 2, \ldots, K$ are used for calculating constituent concentration $\hat{c}_i$ from NIRR data x_{ik}, $k = 1, 2, \ldots, K$ for a new prediction set of samples $i = 1, 2, \ldots, I_{\text{pred}}$.

Prediction of concentration from the spectrum of a new sample is obtained by

$$\hat{c}_i = \hat{g}_0 + \sum_{k=1}^{K} \hat{g}_k \cdot x_{ik} \tag{2}$$

where $\hat{}$ means "statistically estimated."

If the true constituent concentration c_i is known in these new samples, the error in the prediction of the chemical variable, the Standard Error of Prediction (SEP), can be computed:

$$\text{SEP} = \sum_{i=1}^{I_{\text{pred}}} (c_i - \hat{c}_i)^2 / I_{\text{pred}}. \tag{3}$$

This prediction error depends on the complexity of the calibration solution (i.e., on the number of error phenomena modelled by linear factors) as shown in Fig. 3. Increasing the complexity of the calibration may improve the accuracy of prediction by modelling more interferences in the NIRR data. On the other hand, this increasing complexity also increases the sensitivity to random errors in the data. Theoretically, optimal prediction occurs when these two effects (interference reduction and noise sensitivity) balance.

To explore the effects of various NIRR data transformations prior to the linear calibration and prediction the present paper uses the following tools:

1. For data set I (samples with only two chemical compositions, but varying physical structure) NIRR data are plotted against other NIRR data to distinguish between additive and multiplicative spectroscopic effects of physical structure changes. Analysis of variance is used for interpreting the MSC parameters.
2. For data set II (simple mixtures of a lean and a fat sample, each mixture with varying physical structures) spectroscopic data at one wavelength typical for fat is plotted against fat percentage to investigate instrument response linearity.
3. For data set III (real commercial meat samples) multivariate calibration and prediction was performed, and SEP for fat is plotted against model complexity in analogy to Fig. 3, to determine optimal model complexity for the different data transformations.

Linearizing Transformations. *Nonlinear problems.* The spectra from most NIRR instruments are transformed from measured reflectance to apparent absorbance: $A = \log_{10}(1/R)$. In some instruments, including the present one, these A values at different wavelengths $k = 1, 2, \ldots, K$ are used directly as regressors x_{ik} in Eq. 1 and 2 for different samples $i = 1, 2, \ldots, I$.

This transformation is meant to improve the linearity between chemical concentration and spectral data, analogous to classical absorption spectrometry of clear solutions where Beer's law applies:

$$A = 1/T = \sum_{j=1}^{J} \epsilon_j \cdot c_j \cdot l \tag{4}$$

where T is transmission, while ϵ_j and c_j represent extinction coefficient and concentration for chemical constituents no. $j = 1, 2, \ldots, J$ at a given wavelength, and l is the optical path length. For spectral data that follows Eq. 4 the linear multivariate model (Eqs. 1 and 2) applies directly, because Eq. 4 can be linearly transformed to yield c, the concentration of a given constituent (e.g., fat) as:

$$c = \text{linear function of } (A_k, \epsilon_{jk}, j = 1, 2, \ldots, J; \; k = 1, 2, \ldots, K). \tag{5}$$

The assumptions leading to Beer's law are not fulfilled in diffuse spectometry. The apparent absorbance A does not have a completely linear relationship to c for NIRR data, and the linear calibration model (Eqs. 1 and 2) is not fully satisfied. The major source of error is light scatter variation.

The complex physical structures of food samples, with solid particles, liquid droplets, and air inclusions, display widely different refractive indices. Even after thorough blending, samples are unlikely to be optically homogenous. This variability in light scatter blurs the relationship between the spectral data and the actual concentrations of the constituents. Specular reflectance and internal reflection at the inside of the sample's surface also cause linearity problems, with the result that A is not linearly related to c even at constant light scattering levels.

Multivariate linear calibration models do have the ability to approximate certain nonlinear phenomena, but at the expense of an increased calibration model complexity. For instance, *one* nonlinear, "boomerang-shaped" phenomenon appears as *two* linear phenomena spanning the flat plane of the boomerang. Thus, the nonlinearities in A lead to increased statistical model complexity, which again increases the sensitivity to measurement noise and thus increases the prediction error SEP for data from a given set of calibration samples. Some types of nonlinearities may even be expected to defy linear modelling completely.

The following three techniques for improving the linearity of reflectance measurements were tested: (1) The logarithmic transformation from R to A was replaced by the Kubelka-Munk transform. (2) The reflectance data used in these transforms were corrected for internal and external surface reflection according to the Saunderson theory. (3) The light scatter effects were reduced by the Multiplicative Scatter Correction.

Kubelka-Munk Theory. The Kubelka and Munk (K-M) theory,[11] long used in the paint and paper industries, is described fully by Judd and Wyszecki.[12] It is a two-flux radiative transfer theory which takes account of absorption by a coefficient K and scatter by a corresponding coefficient S at a given wavelength. Reflectance R of opaque layers that are optically infinitely thick is termed reflectivity R_∞, and is linearized by conversion to a new unit, termed (K/S):

$$(K/S) = \frac{(1 - R_\infty)^2}{2R_\infty} \tag{6}$$

where K is the absorption coefficient and S the scatter coefficient of the sample at a given wavelength. Given the high light absorption coefficients of biological material in the NIR wavelength range of 1400–2600 nm, R_∞ corresponds to measured reflectance, R.

According to the K-M theory the absorption coefficient K can be written as a linear additive function of chemical constituent concentrations analogous to Beer's law (Eq. 4):

$$K = \sum_{j=1}^{J} k_j \cdot c_j \tag{7}$$

where c_j is the concentration of constituent no. j with specific absorption coefficient k_j.

Provided the scatter coefficient S is constant, the (K/S) data are also linear, additive functions of the concentrations of the chemical constituents:

$$(K/S) = \frac{1}{S} \cdot \sum_{j=1}^{J} k_j \cdot c_j \tag{8}$$

and therefore x_{ik} in Eqs. 1 and 2 can be replaced by $(K/S)_{ik}$. However, if light scatter level varies with homogenization, S will, according to Eq. 6, change the (K/S) data in a *multiplicative* way, not an *additive* way and the (K/S) data cannot be expected to fit well to the linear additive calibration model in Eqs. 1–2. The aim of the MSC is to estimate and eliminate variations in S so that these calibration models can be applied.

Saunderson-Correction for Surface Reflections. A correction to take account of external and internal reflection losses at the boundary between air and the sample surface, introduced by Saunderson,[13] is fully described by Judd and Wyszecki[12] and by Allen.[14] It yields corrected reflectance R^* (Fig. 4):

$$R^* = \frac{R - k_1}{1 - k_1 - k_2 + k_2 R} \tag{9}$$

where k_1 = correction constant for external specular reflection, between 2 and 4%, depending on the refractive index; and

k_2 = correction constant for internal reflection, between 40 and 60% (Allen[14]).

Corrected reflectance R^* can then replace R in the computation of (K/S) data (Eq. 5, 6).

Multiplicative Scatter Correction. The basis of the Multiplicative Scatter Correction (MSC)[4] is the fact that light scatter's wavelength dependency is different from that of chemically based light absorption. Using data from many wavelengths, one can distinguish between absorption and scatter. The scatter for each sample is estimated relative to that of an "ideal" sample, and each sample's spectrum is then corrected so that all samples appear to have the same scatter level as the "ideal."

In this investigation a relatively simple MSC model was used. The assumptions are that for any given sample the scatter coefficient S is the same at all NIR wavelengths and scatter variations due to chemical variations are ignored. As "ideal" sample the average of the calibration set was used. This MSC version is based on the simple linear model for each sample:

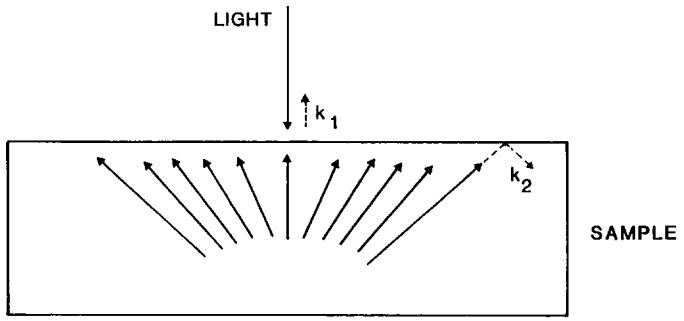

FIG. 4. The Saunderson correction factors for external and internal total reflection at boundary between the sample and its immediate surroundings. Part of the collimated light beam striking the boundary is reflected specularly (k_1); the rest enters the sample and is scattered and partially absorbed by the chemical constituents. The nonabsorbed, diffuse light attempting to emerge back out from the sample is partially reflected back into the sample (k_2) due to the difference in refractive index between sample and its surroundings.

$$X = a + b\bar{X} + E \tag{10}$$

in which X symbolizes the NIRR spectrum of the sample and $\bar{X}$ symbolizes the spectrum of the "ideal" sample, in this case the average spectrum, in some unit (R, R^*, A or K/S). With X and $\bar{X}$ in K-M units, the multiplicative constant b corresponds roughly to $\bar{S}/S$, the inverse of the relative scatter of the sample, and the additive offset term is related to its specular reflectance. The residual spectrum about the MSC model, E, ideally represents the chemical information in X.

For each sample, a and b were estimated by ordinary least-squares regression of spectrum X vs. spectrum $\bar{X}$ over the available wavelengths $k = 1, 2, \ldots, K$, yielding estimates $\hat{a}$, $\hat{b}$, and $\hat{E}$. The corrected spectrum X_1 for each wavelength was calculated by

$$X_1 = (X - \hat{a})/\hat{b} = \bar{X} + \hat{E}/\hat{b}. \tag{11}$$

Figure 5 shows the inter-relationships between the different transformations defined in Table III.

In another paper[15] the effect of eliminating the constant a is discussed.

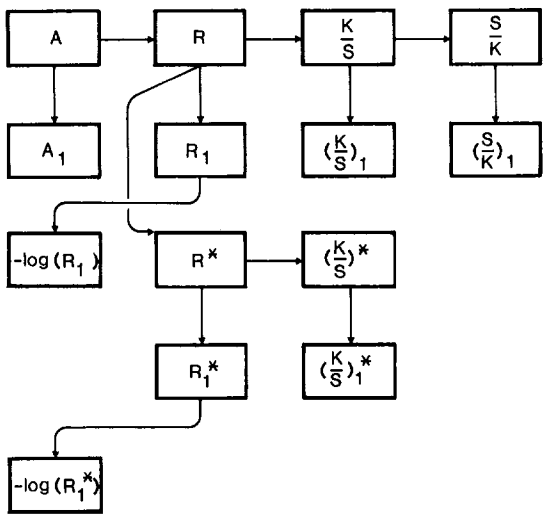

FIG. 5. Schematic view of the nonlinear transformations defined in Table III.

TABLE III. Transformation formulas.

Name	Symbol	Formula	
Reflectance	R	measured	(12)
Absorbance	A	$A = -\log_{10}(R)$	(13)
Kubelka-Munk	K/S or D	$K/S = \frac{(1-R)^2}{2R}$	(14)
Inverse Kubelka-Munk	S/K	$S/K = \left(\frac{K}{S}\right)^{-1}$	(15)
Saunderson	R^*	$R^* = \frac{R - k_1}{1 - k_1 - k_2 + k_2 R}$	(16)
Saunderson + Kubelka-Munk	$\left(\frac{K}{S}\right)^*$	$\left(\frac{K}{S}\right)^* = \frac{(1-R^*)^2}{2R^*}$	(17)
Multiplicative Scatter Correction[a]	X_{1ik}	$X_{1ik} = \frac{X_{ik} - \hat{a}_i}{\hat{b}_i}$	

[a] The letter X represents any of the other symbols: They are given the subscript 1 when MSC-transformed by parameters $\hat{a}_i$ and $\hat{b}_i$, after estimation by simple linear regression, Eqs. 10–11. Values used to test the Saunderson corrections were 0.04 and 0.6 for k_1 and k_2, respectively.

RESULTS AND DISCUSSION

Problems with Scatter and Boundary Reflections Illustrated for Meat Samples of Constant Chemical Composition. Figure 1 shows typical reflectance spectra for low- and high-fat samples, here represented by the spectrum for the 3% (stars) and 17% (filled circles) fat samples in data set 1, averaged over homogenizations and surface qualities.

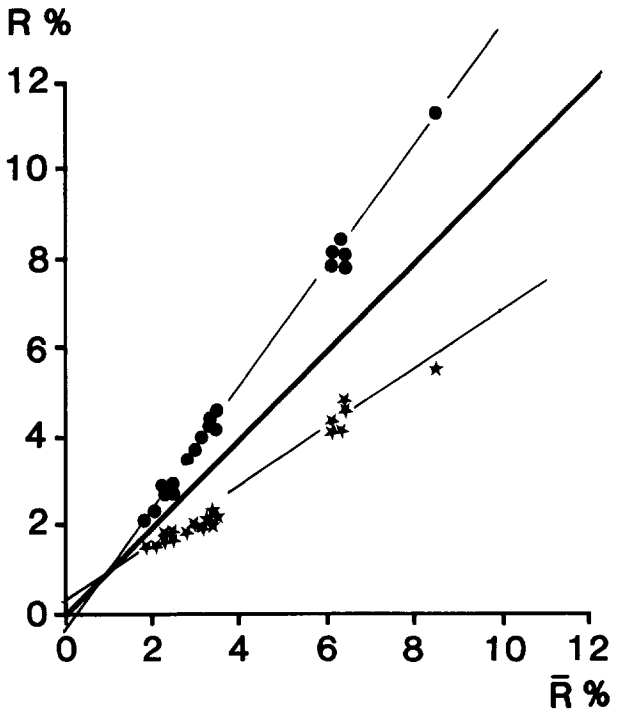

FIG. 6. Spectral changes caused by fat content differences. Abscissa: Average reflectances of all samples in sample set 1. Ordinates: Reflectances of two different fat levels plotted in Fig. 1. Stars: Low-fat sample (3% fat). Circles: High-fat sample (17% fat). Each point represents a wavelength.

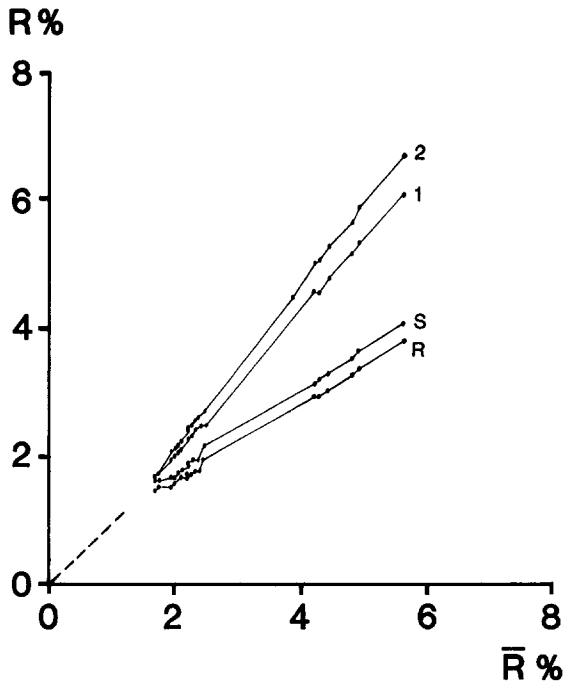

FIG. 7. Reflectances for the 3% fat sample with four different sample treatments, plotted against $\bar{R}$, the average reflectance for the 3% fat sample in data set 1. S = No Homogenization, Smooth surface. R = No Homogenization, Rough surface. 1 = 1 Homogenization, Smooth surface. 2 = 2 Homogenizations, Smooth surface.

The reflectance of the high-fat sample is generally higher than that of the low-fat sample. In addition one can see some minor spectral variations, e.g., between 1700 and 1800 nm, where fat has specific absorption bands. We would like to use this information to calibrate for fat. The problem is that the variations induced in the spectral measurements by variations in sample homogenization and surface quality are similar to the general level change due to fat, and much larger than the small variations at the specific fat absorption bands.

Figure 6 shows the same reflectance spectra as Fig. 1, but plotted in a different way, namely against the total mean reflectance spectrum of sample set 1, instead of against wavelength. In both figures the points represent different wavelengths. In Fig. 6 the thick diagonal line is where the points of a sample would have appeared in case the sample had the same reflectance as the average spectrum. Again it is evident that the low-fat sample has lower reflectance than the high-fat sample. But in addition, this way of plotting the spectra in the wavelength space reveals that these differences can be approximated by straight lines. The small specific variations in the spectra due to chemical light absorptions now appear as minor deviations around these lines.

MSC consists of estimating the intercept (offset a) and slope b of this line for each spectrum by linear regression against the abscissa spectrum. The spectrum is then corrected by a shifting along the ordinate to ensure zero intercept, and rotation of the line till it reaches the diagonal line. Thereby, the corrected spectral data appear to have the same intercept and slope as the average spectrum, while the specific chemical information remains more or less unchanged except for a scaling.

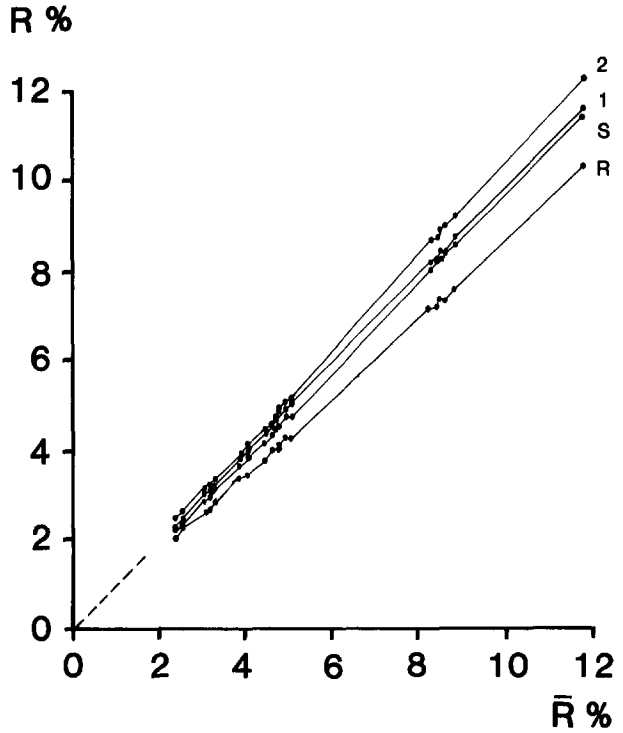

FIG. 8. Reflectances for the 17% fat sample with four different sample treatments, plotted against $\bar{R}$, the average reflectance for the 17% fat sample in data set 1. S = No Homogenization, Smooth surface. R = No Homogenization, Rough surface. 1 = 1 Homogenization, Smooth surface. 2 = 2 Homogenizations, Smooth surface.

In data set 1 the samples were treated with three different degrees of homogenization (0, 1, and 2 times) and two different surface treatments (rough and smooth). Figures 7 and 8 show how the effects of these physical changes affects the reflectances for samples with identical chemical compositions. The two figures represent 3% and 17% fat, respectively, and the curves in each figure show four of the treatments tested in set 1: Smooth (S) and Rough (R) samples with no homogenization, and 1 and 2 times homogenization with smooth surfaces. The four reflectance spectra are here plotted as ordinates against abscissa consisting of the average reflectance spectrum for that fat level (i.e., the spectrum used for ordinate in Figs. 1 and 6). Wavelengths with consecutive reflectance averages have been connected by line segments for visualization. The figure shows that there is a strong effect of both smoothing the surface and homogenizing the samples.

All the reflectance spectra in set 1 were submitted to MSC against the total average reflectance spectrum (the one used as abscissa in Fig. 6). Table IV shows the main results for the intercept a and the slope b from this factorial experiment (2 fat levels × 3 homogenizations × 2 surfaces × 2 replicates).

Compared to the between-replicates variation noise, most of the main effects were large and statistically significant. All the interaction effects were small, most of them hardly larger than the noise. The table shows that the difference in reflectance between fat and lean meat is primarily of multiplicative nature since it appears in b, although some additive offset effect is also found, as shown also in Fig. 6. This fat effect is probably due to

TABLE IV. Effect of homogenization and surface smoothness.[a]

Main effects[b]		b	a
Fat:	3%	0.66	0.19
	17%	1.34	−0.13
Homog.:	No	0.83	0.39
	1 ×	1.05	−0.13
	2 ×	1.13	−0.17
Surface:	Smooth		0.20
	Rough	n.s.	−0.14
Interactions:			
Fat × Homog.		(i)	n.s.
Fat × Surface		n.s.	n.s.
Homog. × Surface		n.s.	(ii)
Fat × Homog. × Surface		n.s.	n.s.
Noise level			
Std. dev. between replicates		0.03	0.07

[a] MSC-parameters a_i and b_i for sample i are estimated from its reflectance spectrum (R_{ik}) from set 1 (i = 2 to 24) by:

$$R_{ik} = a_i + b_i\bar{R}_k + e_{ik} \text{ over wavelengths } k = 1, 2, \ldots, 19$$

where $\bar{R}_k$ is the average spectrum of the samples.

[b] Values for the multiplicative (b) and the additive (a) correction parameters for the three factors were obtained by averaging over the other two factors and over the two replicates. The standard deviation between replicates was used to find nonsignificant (n.s.) effects. All main effects but one are significant at >99%. Two interaction effects were found to be significant, although quite small; these effects are: (1) Extra increase in b from 0 to 1 × homogenization at 3% fat. (2) Largest effect of surface on a for 2 × homogenization, smallest for no homogenization.

the difference in reflective index between fat and muscle tissue, but may also in part be due to differences in the way the two types of meat behave during homogenization, etc.

Increasing the degree of homogenization increases the reflectance systematically: The additive effect a decreases slightly, while the multiplicative effect b increases strongly, especially in lean meat from the lowest to the next degreee of homogenization. This homogenization effect may be a result of the decreased particle size, causing increased light scattering due to the refractive index difference between fat and muscle tissue. The inclusion of air bubbles may also contribute to the increased light scatter.

Smoothing the surface increased the reflectance, and this increase is strongest for well-homogenized samples. This increase is probably a result of increased specular reflectance and decreased local shadowing effects in the surface layer.

Univariate Study of the Linearizing Effect of the Transformations (Data Sets 2 and 3). The purpose of the MSC transformation is the elimination of specular reflectance, scatter errors, etc., in order to linearize the spectral data and decrease noise variance. The effect of this is shown in Fig. 9 for reflectance at 1722 nm, a typical fat absorption peak, against fat content for the mixtures of lean and fat meat in sample set 2. The data before transformation show a fuzzy, nonlinear relationship with positive instead of negative slope. The same data after MSC transformation are much more linear, the noise variance has diminished substantially, and the slope is now negative, as expected from the light absorption of fat.

This MSC transformation was applied to the same

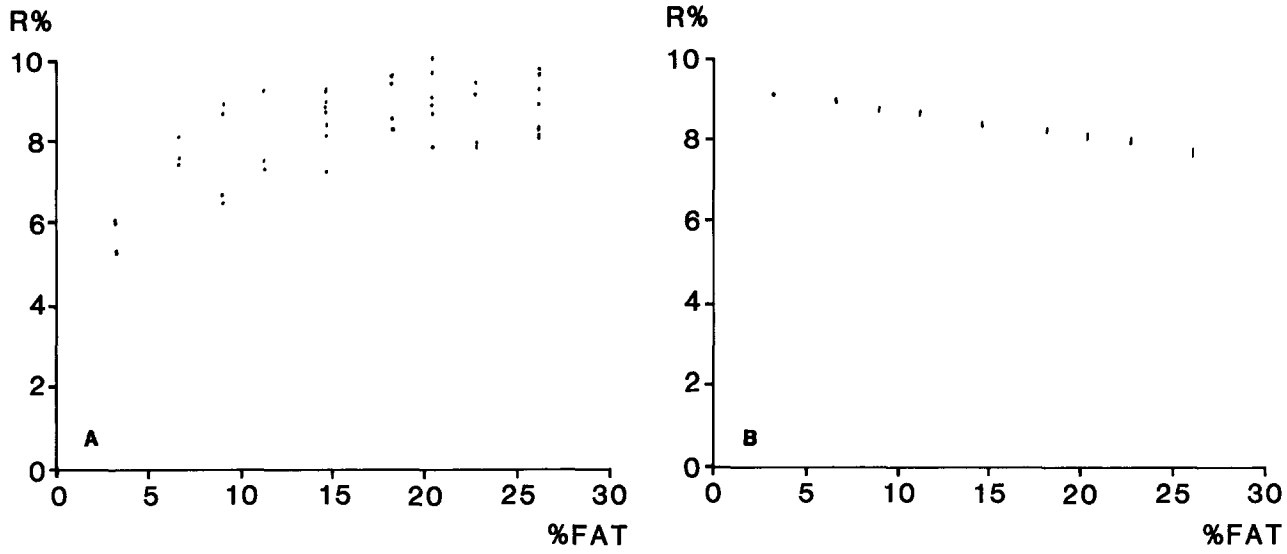

FIG. 9. Reflectance at 1722 nm against fat content for set 2 samples. A, Before MSC transformation. B, After MSC transformation.

NIRR data expressed in A_1, $(K/S)_1$, $(S/K)_1$, R_1* and $\log(1/R_1)$ units, for the same set 2 samples. The $(S/K)_1$ data showed good linearization and small variance, similar to R_1. The A_1 data showed a small variance, but the line showed a somewhat sigmoid shape. The $(K/S)_1$ data still showed both nonlinearity and a large variance.

The Saunderson transformation data R_1* were similar to the reflectance data R_1. In general, the Saunderson transformation had little effect, probably due to the low values of measured reflectances. The $\log(1/R_1)$ results were similar to those of A_1.

Similar results were obtained with the 60 commercial meat samples in set 3. In order that the dependency of the MSC parameters a_i and b_i on the chemical composition of these samples might be studied, they are plotted against fat content in Fig. 10 for reflectance (R). The plots show that the additive distortion term, a_i, shows a slight variation with fat content, being minimum at 15–20 percent fat. The multiplicative distortion term, b_i, increases strongly with fat content of the sample, most drastically from 0 to 20 percent fat. These results correspond well to the main effect of fat observed in set 1 (Table IV). Analogue results were obtained for the other transformations of the NIR data.

Multivariate Modelling to Test the Practical Usefulness of the MSC Transformation (Data Set 3). Because multivariate calibration models are used in practical NIR spectrometry, the effect of linearizing transformations was tested in a multivariate model. The first criterion to be observed is the minimization of the *number of dimensions* needed for calibration, which ideally should not be more than the number of independent chemical components in the system. As discussed earlier, a one-dimensional solution should be almost satisfactory for raw meat, since it consists of muscle and fat tissue summing to 100%.

Another criterion comes from observation of *prediction errors*; the SEP should be as small as possible. The PLS calibration method was tried out for all the transformations in Fig. 5 on data set 3. In Fig. 11 the SEP of

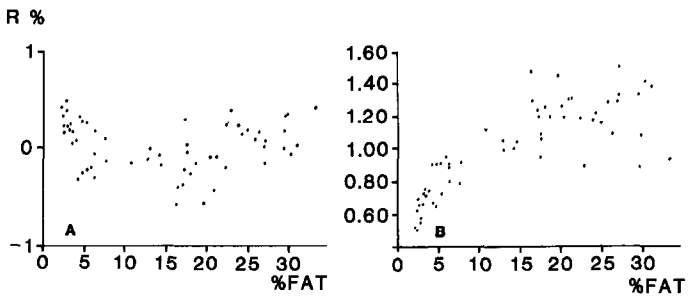

FIG. 10. MSC-correction coefficients a_i (additive) (A) and b_i (multiplicative) (B) for reflectance (R) data from the sample set 3, plotted against fat percentage.

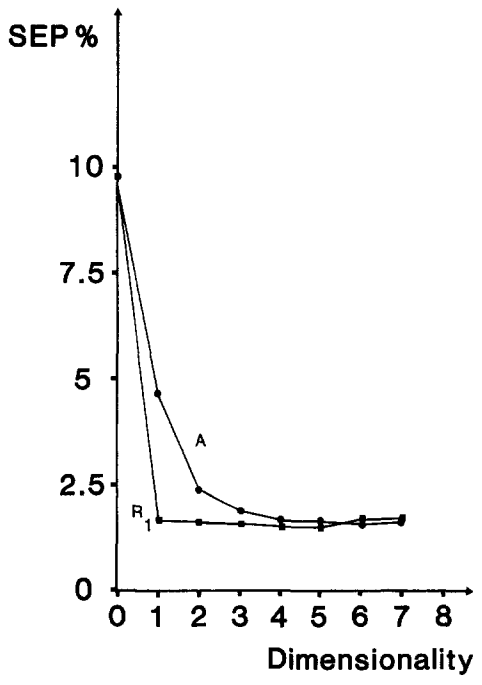

FIG. 11. Prediction error SEP for fat against dimensionality of calibration model, plotted for the conventional unit (A) and for MSC-transformed reflectance (R_1).

TABLE V. SEP for fat using various linearizing transforms for the NIRR data from set 3, and given for increasing model complexity (No. of PLS dimensions).

	0	1	2	3	4	5	6	7	8	9	10
A	9.81	4.61	2.33	1.81	1.58	1.58	1.52	1.58	1.59	1.57	1.58
A_1	9.81	1.86	1.60	1.52	1.48	1.48	1.45	1.47	1.48	1.49	1.53
R	9.81	5.26	2.52	2.18	1.72	1.85	1.72	1.74	1.80	1.83	1.82
R_1	9.81	1.62	1.57	1.52	1.47	1.46	1.60	1.60	1.68	1.70	1.72
$-\log(R_1)$	9.81	1.62	1.57	1.51	1.47	1.46	1.58	1.59	1.62	1.64	1.64
R^*	9.81	5.18	2.51	2.03	1.65	1.74	1.63	1.65	1.71	1.71	1.75
R_1^*	9.81	1.56	1.55	1.53	1.53	1.46	1.58	1.59	1.63	1.68	1.73
$-\log(R_1^*)$	9.81	1.55	1.55	1.52	1.47	1.46	1.58	1.59	1.62	1.68	1.73
$\frac{K}{S}$	9.81	4.43	2.30	1.71	1.69	1.81	1.78	1.72	1.57	1.58	1.61
$\left(\frac{K}{S}\right)_1$	9.81	3.06	1.79	1.49	1.48	1.46	1.46	1.45	1.46	1.47	1.47
$\left(\frac{K}{S}\right)^*$	9.81	4.54	2.48	1.80	1.84	1.87	1.90	1.77	1.62	1.61	1.56
$\left(\frac{K}{S}\right)_1^*$	9.81	3.49	1.99	1.54	1.52	1.48	1.46	1.45	1.47	1.46	1.49
$\frac{S}{K}$	9.81	5.43	2.60	2.61	2.10	2.39	2.20	2.25	2.24	2.30	2.29
$\left(\frac{S}{K}\right)_1$	9.81	1.65	1.58	1.56	1.50	1.50	1.62	1.63	1.70	1.73	1.75

fat is shown against the dimensionality of the calibration model for the conventional unit, apparent absorbance, A, and a MSC unit, R_1. While the A data require 3–4 factors to yield small SEP, the R_1 transformation leads to the expected one-dimensional model.

The SEP for fat is given in Table V for all the NIR transformations as a function of the dimensionality of the calibration model. The two underlined values for each transformation show where about 95% of the original variance in the prediction fat data has been accounted for, and where the minimum is reached, respectively.

The relatively high SEP even with the best transformations is probably due to the wide range of fat concentrations (0–40% fat) covered by these industrial samples and to the less precise sampling and mixing techniques used in their collection. It is difficult to assess statistically the effect of this high prediction error level on the comparison of the different data linearizations and corrections since they are all applied to the same data. The following heuristic interpretation is proposed:

Transformations leading to an almost one-dimensional model, SEP <2.0% after one dimension, are (in order of ascending prediction error after one dimension):

$$-\log(R_1^*),\ R_1^*,\ -\log(R_1),\ R_1,\ (S/K)_1,\ A_1.$$

These transformations appear most useful if a simple calibration is desired, since they require few calibration samples. Transformations leading to a minimal SEP of less than 1.5% are in order of ascending number of dimensions required to reach 1.5%:

$$\left(\frac{K}{S}\right)_1,\ A_1,\ R_1,\ -\log(R_1),\ -\log(R_1^*),\ \left(\frac{K}{S}\right)^*,\ R_1^*.$$

These transformations appear most useful if minimal prediction error is of primary importance. The transformations that satisfy both requirements for this data set are:

$$R_1,\ R_1^*,\ A_1,\ -\log(R_1),\ -\log(R_1^*).$$

CONCLUSION

In the present study the NIR reflectance spectrum of a sample was modelled in terms of *optical* effects causing changes in the direction of the light, and *chemical* effects causing light absorption.

In set 1 it was found that the optical effects in the reflectance spectrum were more or less wavelength-independent in this wavelength region, and could be approximated by an additive offset part, a_i, and a multiplicative proportional part, b_i. The offset was primarily sensitive to surface texture, while the proportionality coefficient was primarily sensitive to sample homogeneity.

A simple regression of each individual sample's spectrum or the average spectrum was used to estimate the offset and the proportionality coefficient. These estimated parameters were then used for standardizing the optical effects prior to the fat determination from the NIR data. This was called "Multiplicative Scatter Correction,"[4] and gave considerably simpler calibration and somewhat improved fat determination results in sets 2 and 3.

Various alternative transformations of the reflectance data, some applied prior to MSC (e.g., $A_1 = [-\log(R)]_1$, some after MSC (e.g., $-\log(R_1)$, were tested to determine their linearizing effects. When MSC was used, all the transformations except those based on Kubelka-Munk theory gave the simple type of calibrations expected for raw meat.

However, even without MSC most of the transformations gave quite good predictive ability (SEP < 2%, i.e., >95% of variance explained) for fat, by the use of more complex calibration models. This shows that even nonlinear relationships can be described by a multidimensional linear model like the PLS regression, as long as the model is made complex enough.

The fat content of the samples was in this paper shown to have both optical effects and chemical light absorption effects. The optical effects of fat on the reflectance spectra were much greater than the chemical effects, but less selective, since the optical effects were also strongly sensitive to sample preparation. They were also more difficult to model by an additive calibration model, due to the multiplicative part, affecting reflectance at all the wavelengths by the same factor. Removal of the optical effects by MSC therefore improved fat prediction.

Future research might include testing the linearity transformations on other data sets and optimizing the Saunderson constants. The MSC parameter b_i could be compared to independent measurements of scatter and particle size. Alternative methods for estimating the parameters a_i and b_i are being investigated.

ACKNOWLEDGMENT

Ingebjørg Pedersen is thanked for good technical assistance. Ragnhild Norang is thanked for skillful typing and Ulla Bråthe for fine graphics. IBM Belgium is thanked for financial support.

1. W. C. Kruggel, R. A. Field, M. L. Riley, H. D. Radloff, and K. M. Horton, J. Assoc. Off. Anal. Chem. **64,** 692 (1981).
2. H. Martens, E. A. Bakker, and K. I. Hildrum, in *Proceedings of 27th European Meeting of Meat Research Workers, Wien, Austria,* O. Prändl, Ed. (1981), pp. 561–564.
3. E. Lanza, J. Fd. Sci. **48,** 471 (1983).
4. H. Martens, S. Å. Jensen, and P. Geladi, *Proceedings, Nordic Symposium on Applied Statistics, Stavanger, June 1983* (Stokkand Forlag Publishers, Skagenkaien 12, Stavanger, Norway, 1983, ISBN 82-90496-02-8), pp. 205–234.
5. K. H. Norris and R. F. Barnes, in *Proceedings, 1st Int. Symp. Feed Composition, Animal Nutrient Requirements and Computerization of Diets* (International Feedstuffs Inst., Utah State University, Logan, Utah, USA, 1977).
6. K. H. Norris, in *Food Research and Data Analysis,* H. Martens and H. Russwurm, Jr., Eds. (Applied Science, London, 1983), pp. 95–114.
7. K. I. Hildrum, M. Valland, and H. Martens, in *Food Research and Data Analysis,* H. Martens and H. Russwurm, Jr., Eds. (Applied Science, London, 1983), p. 416.
8. S. Wold, H. Martens, and H. Wold, in *Matrix Pencils,* A. Ruhe and B. Kågström, Eds. (Springer-Verlag, Heidelberg, 1983), pp. 286–293.
9. H. Martens and S. Jensen, in *Proceedings of the 7th World Cereal and Bread Congress, Prague, June 1982,* F. Holas and F. Kratochvil, Eds. (Elsevier, Amsterdam, 1983), pp. 607–647.
10. H. Martens, O. Vangen, and E. Sandberg, in *Proceedings, Nordic Symposium on Applied Statistics* (Stokkand Forlag, Skagenkaien 12, Stavanger, Norway, 1983, ISBN 82-90496-02-8), pp. 235–268.
11. P. Kubelka and F. Munk, Z. Tech. Phys. **12,** 593 (1931).
12. D. Judd and G. Wyszecki, *Color in Business, Science and Industry* (John Wiley, New York, 1975), 3rd ed., pp. 420–438.
13. J. L. Saunderson, J. Opt. Soc. Am. **32,** 727 (1942).
14. E. Allen, in *AIC Color 77,* F. W. Billmeyer and G. Wyszecki, Eds. (Adam Hilger, Bristol, 1977), pp. 153–179.
15. H. Martens and T. Næs, "Multivariate Calibration by Data Compression," in *Near Infrared Reflectance Spectroscopy,* P. Williams, Ed. (Am. Assoc. Cereal Chem., St. Paul, Minnesota, in press).

Lifetime Measurements of Metastable Levels of Thallium and Lead in the Air-Acetylene Flame by Laser-Enhanced Ionization Spectrometry

N. OMENETTO,* T. BERTHOUD,† P. CAVALLI, and G. ROSSI

Joint Research Center, Chemistry Division, Ispra (Varese), Italy

Laser-enhanced ionization spectrometry with two-step excitation has been used to evaluate the collisional lifetime of the metastable P levels of thallium and lead in an air-acetylene flame burning at atmospheric pressure and supported by a three-slot burner head fitted on a conventional nebulizing chamber. A water-cooled molybdenum electrode immersed in the flame was maintained at high negative potential with respect to the burner body. The ionization current resulting after the two-step excitation was amplified and measured with a digital storage oscilloscope and a boxcar averager. The two excimer lasers were triggered externally with two trigger pulses, one being delayed in time with respect to the other one. In this way the second laser photon, tuned at a transition starting from the metatable level under study, could be correspondingly delayed from the first photon tuned at a transition starting from the ground state.The lifetimes measured were found to be 81 ns and 360 ns for Tl and Pb, respectively.

Index Headings: **Flame spectroscopy; Spectroscopic techniques.**

INTRODUCTION

Laser-enhanced ionization (LEI) or optogalvanic effect in flames at atmospheric pressure is gaining an increasing popularity as a very sensitive technique for trace analysis.[1,2] Several photoexcitation schemes involving one and two laser photons have been used and identified in the literature as "resonance," "nonresonance," and "stepwise" ionization modes of LEI spectrometry. With a two-step excitation scheme, both sensitivity and spectral selectivity are enhanced.[3] Here, the lower level of the transition reached with the second photon is usually the same as the upper level of the transition reached with the first photon. Moreover, both excitation beams are made spatially and temporally coincident in the flame.

Clearly, by delaying in time the second photon with respect to the first one, one can, in principle, obtain the lifetime of the common level. The same reasoning holds even if the excited level reached by the first photon and the starting level of the second transition are not the same, provided that they are strongly collisionally and/or radiatively coupled. This can be the case with elements, such as lead and thallium, with metastable levels lying above the ground state, i.e., levels not radiatively coupled with the ground state. The importance of these levels has been amply recognized in saturation studies

Received 5 August 1984.

* Author to whom correspondence should be sent.
† Present address: DCAEA/SEA/SEACC, Fontenay-aux-Roses, France.

0003-7028/85/3903-0500$2.00/0