Computational Estimation of the Composition of Fat/Oil Mixtures Containing Interesterifications from Gas and Liquid Chromatography Data

Martin H. van Vliet[a], **Geert M.P. van Kempen**[b,*], **Marcel J.T. Reinders**[a], **and Dick de Ridder**[a]

[a]Information and Communication Theory Group, Faculty of Electrical Engineering, Mathematics and Computer Science,
Delft University of Technology, 2628 CD Delft, The Netherlands, and [b]Oil-based Product Technology,
Unilever R&D, Vlaardingen, 3130 AC Vlaardingen, The Netherlands

ABSTRACT: A mathematical framework is introduced that relates analytical data to the composition of fat and oil mixtures. Within this framework, the noise characteristics of four common analytical techniques [FAME, FAME2-pos, CN (carbon number), and AgLC] were investigated and modeled by both additive and multiplicative noise terms. The fat blend recognition (FBR) performance was investigated under these two types of noise, both qualitatively and quantitatively. Furthermore, an extension is proposed that makes it possible to detect interesterifications of unknown mixtures, which was impossible before. The proposed procedure is divided into a qualitative estimation stage, which is focused on identifying the raw materials (RM), followed by a quantitative estimation stage, which is focused on quantifying the levels of the RM identified. We compared two qualitative strategies and four quantitative methods for their ability to correctly estimate simulated mixtures under the noise characteristics determined. The comparison of methods was extended to actual mixtures, revealing promising results. Our analysis presents multiple directions for further adulteration and FBR studies.

Paper no. J11060 in *JAOCS 82*, 707–716 (October 2005).

KEY WORDS: Adulteration, chromatographic data, expectation maximization, fat blend recognition, interesterifications, Lagrange multipliers, oil mixtures, vegetable fats and oils.

The economic advantage of admixing lower-grade products has always been a temptation for commodity suppliers, and various examples of these fraudulent practices can be given (1). Methods for detecting these kinds of malpractices have been the subject of a significant amount of research (2–12). In the case of fat or oil mixtures, it is impossible to physically separate the mixture into its original constituents, which severely complicates detection capabilities. This leaves the problem of detection based on measurable properties of the mixture. Most promising is the TAG (triacylglyceride) distribution of the mixture. The TAG are the main constituents of fats and oils, and the TAG distribution of the mixture is a linear combination of the TAG distributions of the raw materials (RM) (5,13). However, in general the TAG distribution cannot be measured directly: It can only be estimated using a nonlinear one-to-many backward mapping (14). The fact that the relation between the TAG distribution and analytical data is linear explains the frequent use of FAME and carbon number (CN) analysis data (15–17). Moreover, minor components, such as the sterol content, can be used to investigate the composition of fats or oils (9). However, the sterol content can be altered without changing the composition of the mixture itself, thus significantly impairing the ability to detect adulteration.

Research focusing on the detection of adulteration has mainly investigated mixtures of milkfat (MF) and non-milkfat (non-MF) as well as cocoa butter (CB) and cocoa butter equivalents (CBE). Precht (2) investigated linear models based on CN data capable of detecting non-MF additions to MF samples. Ulberth (3) studied the same problem and investigated the applicability of a classification approach based on CN data. Ulberth (4) also considered the problem with only a single non-MF addition and proposed a method based on partial least squares optimization of the CN data. Simoneau (5) proposed identifying the presence of CBE in CB samples by means of scatterplots of a few TAG. Dionisi *et al.* (6) studied both linear and quadratic models based on TAG and TAG ratios to determine the CBE content in CB samples.

A more general approach to detecting adulteration is to focus not only on determining addition levels but also on quantifying all RM used. This is often referred to as fat blend recognition (FBR) and could serve applications in the domain of patent infringement detection.

Van Niekerk and Burger (7) investigated mixtures of, at most, seven RM and proposed a least-squares approach based on FAME data. Van Niekerk and Hasty (8) proposed a few optimizations of this method. Abuhadeed and Kotb (9) used the unconstrained pseudoinverse to quantify mixtures based on FAME data. De Jong and De Jong (10) investigated a brute force method to handle a situation in which many RM are to be considered, also using FAME data. De Jong and De Jong also considered a mixed integer programming branch-and-bound method.

A special problem is posed by mixtures containing chemically processed RM. One frequently applied chemical processing technique is interesterification (13). None of the investigations mentioned above provide any means for detecting interesterifications.

---

*To whom correspondence should be addressed at Oil-based Product Technology, Unilever R&D Vlaardingen, Olivier van Noortlaan 120, P.O. Box 114, 3130 AC Vlaardingen, The Netherlands.
E-mail: Geert-van.Kempen@Unilever.com

In this paper we present an investigation of the noise characteristics of four analytical techniques and propose a noise model. Both qualitative and a quantitative approaches to FBR are presented. Furthermore, a novel strategy is proposed for detecting interesterifications of unknown mixtures. The approaches are compared based on simulated data and real data.

THEORETICAL BACKGROUND

Mathematical problem formulation. Each fat can be uniquely described by its TAG distribution, which can be represented as a vector **x**, and its analytical data can be grouped in a vector **y** (14). The space containing all possible TAG distributions can be referred to as *X*, and the space containing all possible combinations of analytical data as *Y*. A mixture of fats and oils can then be represented as a vector **w**, which corresponds to a point in both *X* and *Y*:

$$\mathbf{x} = \mathrm{M}_x\mathbf{w} \tag{1}$$

$$\mathbf{y} = \mathrm{M}_y\mathbf{w} \tag{2}$$

where M_X represents a matrix in which each column represents the TAG distribution of one of the RM and M_Y represents a matrix in which each column represents the analytical data of one of the RM. The **w** is the same in both equations.

Figure 1 shows the spaces and their interrelations. FBR aims to uncover the vector **w** by obtaining an estimated vector $\mathbf{w}_e$ that resembles the real vector $\mathbf{w}_r$ as closely as possible:

$$\varepsilon_w = d(\mathbf{w}_r, \mathbf{w}_e) \tag{3}$$

where *d* indicates a distance measure between its two arguments. The main difficulty in minimizing Equation 3 is the fact that $\mathbf{w}_r$ is unknown. However, using Equation 1, we can also define a distance measure in terms of the TAG distributions:

$$\varepsilon_x = d(\mathbf{x}_r, \mathrm{M}_x\mathbf{w}_e) \tag{4}$$

Since the TAG distribution of a mixture can be estimated only by a nonlinear one-to-many mapping (14), we make use of the analytical data and focus on defining a distance measure in the analytical space:

$$\varepsilon_y = d(\mathbf{y}_r, \mathrm{M}_y\mathbf{w}_e) \tag{5}$$

One of the main considerations for choosing a distance measure is the type of noise present. This will be investigated in detail in the following paragraphs.

Physical properties of mixtures impose certain constraints on $\mathbf{w}_e$, namely:

$$w_{e,i} \geq 0 \;\forall\; i \in \{1 \ldots n_{\mathrm{RM}}\} \tag{6}$$

$$\sum_{i=1}^{n_{RM}} w_{e,i} = 1 \tag{7}$$

where n_{RM} indicates the number of RM that is being considered and $w_{e,i}$ are the individual weight values.

In addition to these constraints, there are extra constraints from expert FBR knowledge:

$$w_{e,i} \;\geq\; l_i \quad \forall i \in \{1 \ldots n_{\mathrm{RM}}\} \tag{8}$$

$$\sum_{i=1}^{n_{\mathrm{RM}}} I(w_{r,i} > 0) \leq k \text{ with } k \in \mathbb{N}, k << n_{\mathrm{RM}} \tag{9}$$

where *I* is the indicator function, l_i defines a lower limit for the fractions, and *k* is a scalar. The constraint in Equation 8 imposes a lower limit l_i on each of the weight fractions. For instance, knowledge about the quantity of certain fractions may be available. The constraint in Equation 9 states that the number of RM used in a mixture is usually limited. It is highly unlikely, for example, to find more than six RM in a commercial fat mixture.

The FBR strategy we propose consists of two parts: a quantitative estimation and a qualitative estimation. They are discussed separately in the following sections.

Noise characteristics of the analytical data. The analytical techniques will introduce different noise characteristics. We assume the following noise model for analytical noise:

$$\mathbf{y}_n = \mathbf{y}_r + c_1\mathbf{y}_r\boldsymbol{\eta}_1 + c_2\boldsymbol{\eta}_2 \tag{10}$$

where $\mathbf{y}_n$ is the noisy version of $\mathbf{y}_r$, $\boldsymbol{\eta}_i$ represent independent vectors of Gaussian distributed variables with mean 0 and standard deviation 1, and c_1 and c_2 are noise parameters that indicate the noise strength. More precisely, c_1 determines the amount of multiplicative noise and c_2 the additive noise. To impose nonnegativity and summation to one, the resulting $\mathbf{y}_n$ needs to be clipped and normalized. For each analytical technique, the appropriate noise parameters can be estimated from

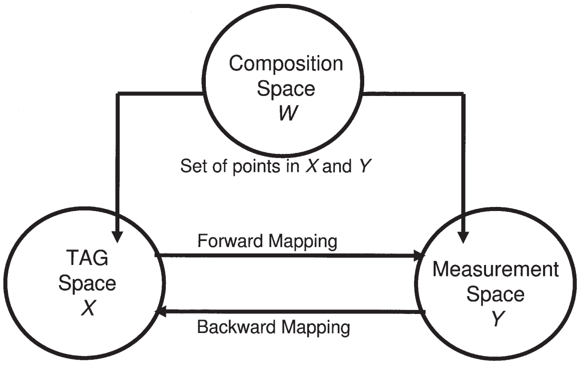

FIG. 1. Schematic representation of the TAG space *X*, measurement space *Y*, and composition space *W*, including their interrelations. Each $\mathbf{w} \in W$ corresponds to a vector in both *X* and *Y*. The forward mapping indicated is a many-to-one linear matrix multiplication, whereas the backward mapping is best described by a nonlinear one-to-many mapping (14).

a collection of test samples, as will be shown in the Experiments section.

In real applications, there will also be noise in the form of natural variance. The characteristics of and typical ranges for this variance are unknown, and are therefore excluded from the model.

Quantitative FBR. The RM in a mixture can be quantified by choosing a distance measure in Equation 5 and minimizing ε_y with respect to $\mathbf{w}_e$. For correct quantitative estimation, it is imperative to include all the appropriate RM in $\mathbf{M}_y$, as it is impossible to detect an RM that has not been included. Based on the known $\mathbf{y}_n$ and $\mathbf{M}_y$ parameters, $\mathbf{w}_e$ can be found. The following sections present an investigation of two possible distance measures.

(i) Differential distance measure. Minimization of ε_y based on a differential distance measure can be formulated as a minimization of the weighted mean squared error (MSE):

$$\varepsilon_y = \arg \min_{w_e} \frac{1}{n_{\mathrm{RM}}} \left\| \mathrm{diag}(\mathbf{a})(\mathbf{y}_n - \mathbf{M}_y \mathbf{w}_e) \right\|^2 \qquad [11]$$

subject to

$$w_{e,i} \geq 0 \quad \forall i \in \{1..n_{\mathrm{RM}}\}$$

$$\sum_{i=1}^{n_{\mathrm{RM}}} w_{e,i} = 1$$

where diag($\mathbf{a}$) is a matrix with the weight elements of $\mathbf{a}$ on the diagonal and $\|\mathbf{x}\|$ represents the 2-norm of vector $\mathbf{x}$. We chose to use the regular MSE with $\mathbf{a} = 1$ since optimization with respect to $\mathbf{a}$ is highly problem specific. Considerations for preferentially matching individual experimental measurements could be the knowledge that certain measurements have a higher spread due to natural variation.

A mathematically consistent method for solving the given constrained minimization problem of Equation 11 is using Lagrange multipliers (18,19). A solution using Lagrange multipliers with only inequality constraints is provided by the nonnegative least squares (NNLS) algorithm (20). However, this algorithm cannot be straightforwardly extended to include the equality constraint as well. By normalizing the solution, we can create a suboptimal solution. The solution based on the normalized NNLS algorithm will be referred to as "L1". We have also derived a Lagrange multiplier solution that can handle both types of constraints by iteratively applying the Lagrange multiplier solutions to the two separate constraint classes until convergence is reached. This method, denoted as "L2," ensures a solution that is optimal in terms of the optimization criterion. The algorithm in pseudocode is listed in Figure 2.

```
Lagrange Method Main Code:
$\mathbf{w}_{e,k=0} = \mathbf{M}_y^{\dagger}\mathbf{y}_n$
$\mathbf{w}_{e,k=1} =$ Lagrange Equality$(\mathbf{w}_{e,k=0})$
While any $\mathbf{X}\mathbf{w} < 0$
    $\mathbf{w}_{e,k+1} =$ Lagrange Inequality$(\mathbf{w}_{e,k})$
    $\mathbf{w}_{e,k+2} =$ Lagrange Equality$(\mathbf{w}_{e,k+1})$
end While


Lagrange Equality$(\mathbf{w}_{e,k})$
$\lambda = \frac{1-\sum(\mathbf{w}_{e,k})}{\sum\left[(\mathbf{M}_y^T\mathbf{M}_y)^{-1}I\right]}$
$\mathbf{w}_{e,k+1} = \left(\mathbf{M}_y^T\mathbf{M}_y\right)^{-1} I\lambda + \mathbf{w}_{e,k}$


Lagrange Inequality$(\mathbf{w}_{e,k})$
Initialization:
$\mathbf{w}_{e,u} = \mathbf{w}_{e,k}$
$V_a = \{\}$
$V_p = \{1 \ldots n_{\mathrm{RM}}\}$
$\mathbf{S} = \mathbf{X}\left(\mathbf{M}_y^T\mathbf{M}_y\right)^{-1}\mathbf{X}^T$
Flag = 1
$\mu = \mathbf{0}$
$\mu' = \{\}$
Main Loop:
While any $\mathbf{X}\mathbf{w}_{e,c} < 0$
    Inner Loop:
    While any $\mu < 0$ or Flag = 1
    Flag = 0
    $\mathbf{P} = \mathbf{S}_{i,i}\ i \in V_a$
    $\mathbf{w}_{e,u,a} = \mathbf{w}_{e,u,i}\ i \in V_a$
    $\mu_a = -\mathbf{P}^{-1}\mathbf{w}_{e,u,a}$
    $\mu_i = 0\ i \in V_p$
    $\mu_i = \mu_a\ i \in V_a$
    n = sum$(\mu' = \mu)$
    $(\max - n)^{\mathrm{th}}$ component from $V_a$ to $V_p$
    $\mu' = \{\mu', \mu\}$
    End Inner Loop:
Flag = 1
$\mathbf{w}_{e,c} = \left(\mathbf{M}_y^T\mathbf{M}_y\right)^{-1}\mathbf{X}^T\mu + \mathbf{w}_{e,u}$
if $\min(\mathbf{X}\mathbf{w}_{e,c}) < 0$
    Move that component from $V_p$ to $V_a$
end if
End Main Loop
$\mathbf{w}_{e,k} = \mathbf{w}_{e,c}$
```

FIG. 2. Pseudocode for the Lagrange methods "L2" and "L3." The top lists the main code (with k being the iteration), which iteratively calls the two subroutines: Lagrange Equality and Lagrange Inequality. For the L2 method, $\mathbf{X}$ is the unity matrix, and for the L3 method it is defined by Equation 21 (see Theoretical Background section).

(ii) Relative distance measure. Csiszár (21) has postulated a set of axioms (consistency, distinctness, continuity, locality, and composition consistency) that a relative distance measure should satisfy. Although it is beyond the scope of this paper to discuss these axioms in detail, the conclusion that can be drawn from them is that for real valued functions having both positive and negative values, only the MSE is consistent with these axioms. For functions that must be nonnegative, the information-divergence (I-divergence) is the only selection rule that satisfies all five axioms. The I-divergence is defined as:

$$\varepsilon_y = \mathbf{y}_n^T \log\left(\frac{\mathbf{y}_n}{\mathbf{M}_y w_e}\right) - \sum(\mathbf{y}_n - \mathbf{M}_y \mathbf{w}_e), \qquad [12]$$

where the division of vectors is done element-wise. It can be shown that the I-divergence is minimized by maximizing the log-likelihood function of a Poisson distributed variable (Poisson noise is multiplicative noise) (22). Such a log-likelihood function can be maximized with the iterative expectation maximization algorithm (EM) (23). In the remainder of this paper, this method will be referred to as EM.

Qualitative FBR. The expert knowledge that a mixture will have a limited number of RM is mathematically described by Equation 9. It is difficult to extract this qualitative information from the available data for a mixture of unknown origin and will vary from case to case, making mathematical treatment problematic. Moreover, the quantitative approach has the mathematical property that adding an extra RM will always result in the same or a decreased ε_y. This is because the result of Equation 11 decreases when the number of RM considered is enlarged. The opposite, leaving out an RM, will result in the same or a higher ε_y.

To find the correct set of RM present in the mixture, Van Niekerk and Burger (7) proposed using backward elimination, leaving out the smallest fraction in $\mathbf{w}_e$, to obtain a solution with no fractions smaller than a preset threshold. This method is referred to as "Niekerk" in the remainder of the paper. Alternatively, one can observe the change in ε_y. We propose a qualitative estimation that stops when ε_y increases more than a certain threshold. This method will be referred to as "Vliet."

Detection of interesterifications of unknown mixtures. In reality, FBR is problematic because the initial RM can undergo chemical processing, which changes the constituents. One frequently applied chemical processing technique is interesterification (13), denoted as (for two fats, F_a and F_b):

$$\mathrm{in}[\alpha F_a/(1-\alpha)F_b] \qquad [13]$$

where α is a parameter that defines the ratio of the two fats ($0 \leq \alpha \leq 1$). Because TAG have only three FA attached to them, it is possible to create all interesterifications in the entire range of α with four points along the trajectory described by α. The set of four points is denoted as preset in(RM).

This can be proven as follows: The FA distributions $F(F_a)$ and $F(F_b)$ are constants that are scaled by α and $1-\alpha$, respectively, and summed to give the FA distribution of the interesterification:

$$F\{\mathrm{in}[\alpha F_a/(1-\alpha)F_b]\} = \alpha F(F_a) + (1-\alpha)F(F_b) \qquad [14]$$

The TAG distribution of the interesterification can be obtained by multiplying the respective FA distribution values of the combined FA distribution for each TAG. After interesterification, FA on the TAG originate from either F_a, denoted as FA_a, or F_b, denoted as FA_b. Given that an FA can originate from two fats, a total of eight combinations can be formed:

$$\begin{array}{l}
sn-1 \quad sn-2 \quad sn-3 \\
\mathrm{FA}_a * \mathrm{FA}_a * \mathrm{FA}_a * \alpha^3 \\
\hline
\mathrm{FA}_a * \mathrm{FA}_a * \mathrm{FA}_b * \alpha^2(1-\alpha) \\
\mathrm{FA}_a * \mathrm{FA}_b * \mathrm{FA}_a * \alpha^2(1-\alpha) \\
\mathrm{FA}_b * \mathrm{FA}_a * \mathrm{FA}_a * \alpha^2(1-\alpha) \\
\hline
\mathrm{FA}_a * \mathrm{FA}_b * \mathrm{FA}_b * \alpha(1-\alpha)^2 \\
\mathrm{FA}_b * \mathrm{FA}_a * \mathrm{FA}_b * \alpha(1-\alpha)^2 \\
\mathrm{FA}_b * \mathrm{FA}_b * \mathrm{FA}_a * \alpha(1-\alpha)^2 \\
\hline
\mathrm{FA}_b * \mathrm{FA}_b * \mathrm{FA}_b * (1-\alpha)^3
\end{array} \qquad [15]$$

To obtain the amount for a TAG after interesterification, one should substitute the respective FA values and sum them. Since this results in a (polynomial) function of α with only four factors (third order), the TAG distribution of an interesterified mixture of F_a and F_b can be constructed by selecting four points from the entire range of α. This is possible since $F(F_a)$ and $F(F_b)$ are constants.

Similarly, it can be shown that for interesterifications of more than two RM (n_{RM}), the number of preset in(RM) [$n_{\mathrm{in(RM)}}$] required for perfect reconstruction is:

$$n_{\mathrm{in(RMs)}} = \frac{n_{\mathrm{RM}}^3 + 3n_{\mathrm{RM}}^2 + 2n_{\mathrm{RM}}}{6} \qquad [16]$$

which is equivalent to the formula indicating the number of TAG that can be formed based on the number of FA, excluding isomers (15).

The coefficients for perfect reconstruction along the entire range of α can be calculated as follows:

$$\mathbf{w}(\alpha) = \mathbf{M}_x^{\dagger}\mathbf{x}(\alpha) \qquad [17]$$

where $\mathbf{M}_x$ is the matrix with the TAG distributions of the preset in(RM) as columns, $\mathbf{M}_x^{\dagger}$ denotes the Moore–Penrose pseudoinverse of $\mathbf{M}_x$ (24), and $\mathbf{x}(\alpha)$ denotes the TAG distribution of the interesterification as a function of α. Figure 3 shows an example of $\mathbf{w}(\alpha)$ for all interesterifications of a mixture containing palm oil (PO) and palm kernel oil (PK).

Note that for certain $\mathbf{x}(\alpha)$ the coefficients in $\mathbf{w}(\alpha)$ are negative. Allowing negative fractions for the preset in(RM) in the Lagrange framework requires thoughtful setting of the lower limits. However, lower limits smaller than zero require the enforcement of two additional inequality constraints:

$$\beta = \sum_{i \in in(\mathrm{RM}s)} w_i \ \geq \ l_\beta \qquad [18]$$

$$\sum_{i \in in(\mathrm{RM}s)} m_{y,j,i} w_i \ \geq \ 0 \ \ \forall j \in \{1..n_y\} \qquad [19]$$

where $\mathbf{M}_x$ indicates the amount of interesterification present, l_β denotes a corresponding lower limit, $m_{y,j,i}$ denotes the element in $\mathbf{M}_y$ at row j and column i, and n_y denotes the number of ele-

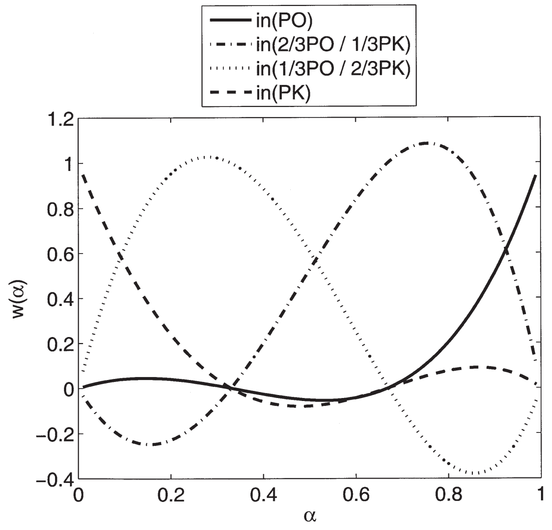

FIG. 3. A plot of $\mathbf{w}(\alpha)$ for the combination of palm oil (PO) and palm kernel oil (PK). The curves presented indicate the amounts required for perfect reconstruction, as defined by Equation 17 (see Theoretical Background section). The curves are based on four equidistantly chosen in(RM) (see the *Detection of interesterifications of unknown mixtures* section) and have been plotted as a function of α.

ments in $\mathbf{y}_e$. Equation 19 ensures that the combined contribution of the preset in(RM) to $\mathbf{y}_e$ is positive. For ease of further derivation, Equation 18 is rewritten into:

$$w_{e,i} - \beta l_i \geq 0 \; i \in \text{in(RM)} \qquad [20]$$

which implies scaling of the lower limits by the amount of interesterification present.

We propose to adapt the Lagrange framework such that the weights representing the in(RM*)* components may become negative. Furthermore, we propose to extend the framework with the additional inequality constraints by combining the in-

$$\mathbf{X} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \ddots & 0 & \vdots & \ddots & \vdots \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1-l_i & -l_i & -l_i \\ \vdots & \ddots & \vdots & -l_i & 1-l_i & -l_i \\ 0 & \cdots & 0 & -l_i & -l_i & 1-l_i \\ 0 & \cdots & 0 & 1 & \cdots & 1 \\ 0 & \cdots & 0 & & & \\ \vdots & & \vdots & & & \\ \vdots & & \vdots & & M_y & \\ \vdots & & \vdots & & & \\ 0 & & \cdots & 0 & & \end{bmatrix} \begin{matrix} (6) \\ \\ \\ \\ (20) \\ \\ (18) \\ \\ \\ (19) \\ \\ \\ \end{matrix} \qquad [21]$$

where the first columns correspond to the regular RM and the last columns to the added preset in(RM). This changes the inequality constraint (Eq. 6) into:

$$\mathbf{X}\mathbf{w}_e \geq 0 \qquad [22]$$

This adapted Lagrange method will be referred to as "L3" in the succeeding sections, and the pseudocode is listed in Figure 2. Extension of the EM algorithm with these constraints is impossible because of the nonnegativity requirement of $\mathbf{w}_e$.

Furthermore, an extra step is proposed in the qualitative estimation procedure to detect the combination of F_a and F_b in the interesterification. The ε_y is calculated for the group of RM and for every combination of pairs of RM that can be in the interesterification. This implies a brute-force search for the pair of RM giving the lowest ε_y, which is selected for further processing in the backward elimination procedure.

A tricky step is converting the estimated in(RM), obtained after the quantitative estimation, into a single interesterification. Using Equation 18, we determine the amount of interesterification present. Uncovering the α is a bit more tedious and can be achieved by minimizing:

$$\varepsilon_x = \underset{\alpha}{\arg\min} \; \| \mathbf{x}(\alpha) - \mathbf{M}_x \mathbf{w}_e \|^2 \qquad [23]$$

where $\mathbf{M}_x$ is a matrix with the TAG distributions of the in(RM) as columns and $\mathbf{w}_e$ is a vector containing the in(RM) fractions. Using Equations 18 and 23, we convert the in(RM) fractions into α and β, giving a converted vector denoted $\mathbf{w}_{e,c}$. This is schematically indicated in Figure 4.

EXPERIMENTS

Both the qualitative and the quantitative methods were tested on simulated data by the following procedure:

1. Select a set of RM.
2. Select the analytical techniques.
3. Construct $\mathbf{M}_y$.
4. Generate a random $\mathbf{w}_r$ and related $\mathbf{w}_{r,c}$.
5. Calculate $\mathbf{y}_r$, and generate $\mathbf{y}_n$.
6. Estimate $\mathbf{w}_e$ based on $\mathbf{M}_y$ and $\mathbf{y}_n$ and convert it into $\mathbf{w}_{e,c}$.
7. Compare $\mathbf{w}_{e,c}$ with $\mathbf{w}_{r,c}$. [24]

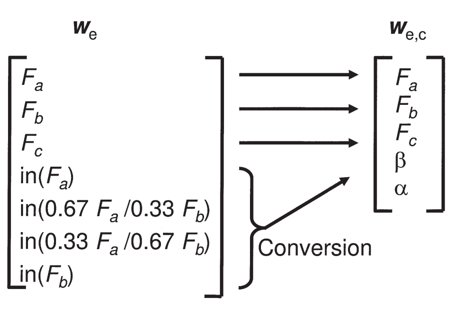

FIG. 4. Schematic indication of the process of estimating the preset in(RM) [along with the regular raw materials (RM)] and converting it into α and β. In the case presented here, three regular fats, i.e., F_a, F_b, and F_c, are considered and an interesterification (in) of F_a and F_b (not necessarily the same as the regular RM). The vector is first labeled $\mathbf{w}_e$ and after conversion $\mathbf{w}_{e,c}$ (see Theoretical Background section). See Figure 3 for other abbreviation.

To compare the real $\mathbf{w}_{r,e}$ after conversion with the estimated $\mathbf{w}_{e,c}$ we used:

$$\begin{aligned}\varepsilon_w = {} & [1-\mathbf{w}_{r,c}(\beta)]\left\|\mathbf{w}_{r,c}(\mathrm{RM},\beta)-\mathbf{w}_{e,c}(\mathrm{RM},\beta)\right\|^2 \\ & +\mathbf{w}_{r,c}(\beta)\left\|\mathbf{w}_{r,c}(\alpha)-\mathbf{w}_{e,c}(\alpha)\right\|^2\end{aligned} \qquad [25]$$

where ε_w signifies the distance measure that defines how well the estimated $\mathbf{w}_{e,c}$ matches the generated $\mathbf{w}_{r,c}$. Note that if the qualitative estimation of the combination of RM in the interesterification is incorrect, one compares two β values corresponding to a different RM combination. This was penalized by inserting a zero in both vectors, indicating that the β values are compared to zero.

The simulations were performed with a set of eight RM. This set consisted of three oils, i.e., bean oil (BO), sunflower oil (SF), and rapeseed oil (RP); two medium-melting fats (MMF), i.e., hardened coconut oil (CN31) and PK; and three high-melting fats (HMF), i.e., PO, palm stearin (dfPOs52mp), and multifractionated palm stearin (mfPOs). Furthermore, all interesterifications based on a mixture of an arbitrary HMF and an arbitrary MMF or oil were considered (thus, there were 15 possible combinations). All $\mathbf{w}_r$ vectors were qualitatively generated by randomly selecting one oil/MMF, one HMF, one random interesterification, and a second HMF (80% chance) or oil/MMF (20% chance). Hence, all vectors had four different RM (counting the interesterification as a single RM). Knowledge about this *a priori* distribution was explicitly excluded from the estimates made so as to obtain unbiased estimates.

The quantitative amounts were generated so as to resemble realistic mixtures by:

$$\begin{array}{ll} \multicolumn{2}{c}{80\%\text{ chance 2 of HMF}} \\ \text{oil/MMF}_1: & \gamma_1 \\ \beta: & (1\text{-}\gamma_1)\gamma_2 \\ \alpha: & \gamma_4 \\ \text{HMF}_1: & (1-\gamma_1)(1-\gamma_2)\gamma_3 \\ \text{HMF}_2: & (1-\gamma_1)(1-\gamma_2)(1-\gamma_3) \\ \multicolumn{2}{c}{20\%\text{ chance 2 of oils/MMF}} \\ \text{oil/MMF}_1: & \gamma_1\gamma_3 \\ \text{oil/MMF}_2: & \gamma_1(1-\gamma_3) \\ \beta: & (1\text{-}\gamma_1)\gamma_2 \\ \alpha: & \gamma_4 \\ \text{HMF}_1: & (1-\gamma_1)(1-\gamma_2) \end{array} \qquad [26]$$

$\gamma_{1,2,3}$ are random Gaussian distributed variables with mean 0.5 and standard deviation 0.1, and the interesterification variable γ_4 is a random Gaussian distributed variable with mean 0.5 and standard deviation 0.2. Nonnegativity and summation to 1 were incorporated by clipping and normalizing the resulting $\mathbf{w}_r$. In the simulations FAME, FAME2-pos, CN, and AgLC, analytical data were used, taken from an internal database.

The real mixtures analyzed were made using commercially available products obtained from Unilever's factory in Rotterdam (The Netherlands).

All calculations were performed in the MATLAB 7 programming environment (25).

Noise characteristics of the analytical data. It is imperative to have a correct understanding of the noise present in the analytical data to develop a suitable method for FBR. Therefore, we set up an experiment to determine the noise parameters of the four analytical techniques used throughout this paper (FAME, FAME2-pos, CN, and AgLC). For each of the four methods, a reference sample was analyzed over a period of a year. For the FAME analysis (ISO 5508 and 5509), 24 measurements of reference sample bcr-164 (26) were performed. Similarly, for the FAME2-pos [IUPAC method 2.210 (27)], 18 measurements of sample in(PO) were taken. For the CN, (ISO CD 22508) 40 measurements of sample 60PK39/40RP70 were taken, and for the AgLC (as described in Ref. 28), 26 measurements of sample in(PO) were performed. For each of the methods, we studied the relation between the means and standard deviations. The coefficients c_1 and c_2 in the proposed noise model (Eq. 10) were fitted by trial and error, matching the regression lines. The resulting means and standard deviations are shown in Figure 5. The c_1 and c_2 values determined are listed in Table 1.

As can be seen from Figure 5 and Table 1, the FAME, FAME2-pos, and CN methods had both an additive and a multiplicative noise component, whereas the AgLC method showed only an additive term. These noise coefficients were used in all simulation experiments throughout this paper.

RESULTS AND DISCUSSION

Qualitative FBR. To test the qualitative estimation procedure, we generated 2,000 random $\mathbf{w}_r$ vectors as described, with four nonzero RM. Figure 6 shows both $\overline{\varepsilon_y}$ and $\overline{\varepsilon_w}$ as a function of the number of RM. The plot of the $\overline{\varepsilon_y}$ shows a descending line as expected. The plot of $\overline{\varepsilon_w}$ shows a dip at four RM, which is to be expected because the original generated $\mathbf{w}_r$ had four nonzero RM. This clearly shows the advantage that can be gained by correct qualitative estimation of a mixture. Simultaneous with the dip in $\overline{\varepsilon_w}$, one can discern a bend in the $\overline{\varepsilon_y}$ plot, justifying a strategy focused on detecting this.

Both the Niekerk and Vliet methods provide a stopping criterion for the backward elimination procedure. To compare them and investigate their performance, we generated a set of 1,000 different random $\mathbf{w}_r$ vectors and, based thereon, 1,000 $\mathbf{y}_n$. Using the L1, L2, L3, and EM methods, we performed qualitative estimations, and the resulting $\mathbf{w}_e$ were compared qualitatively to $\mathbf{w}_r$. We varied the Vliet threshold from 0 to 100% in steps of 1% and the Niekerk threshold from 0 to 0.10 in steps of 0.001. The number of true positives (TP) as a function of the thresholds was then derived. A TP is defined as a $\mathbf{w}_e$ for which the complete set of RM has been correctly identified, including a correct estimation of the combination of RM in the interesterification.

Figure 7 shows the result of the 1,000 qualitative estimations as a function of the threshold levels. The figure shows a

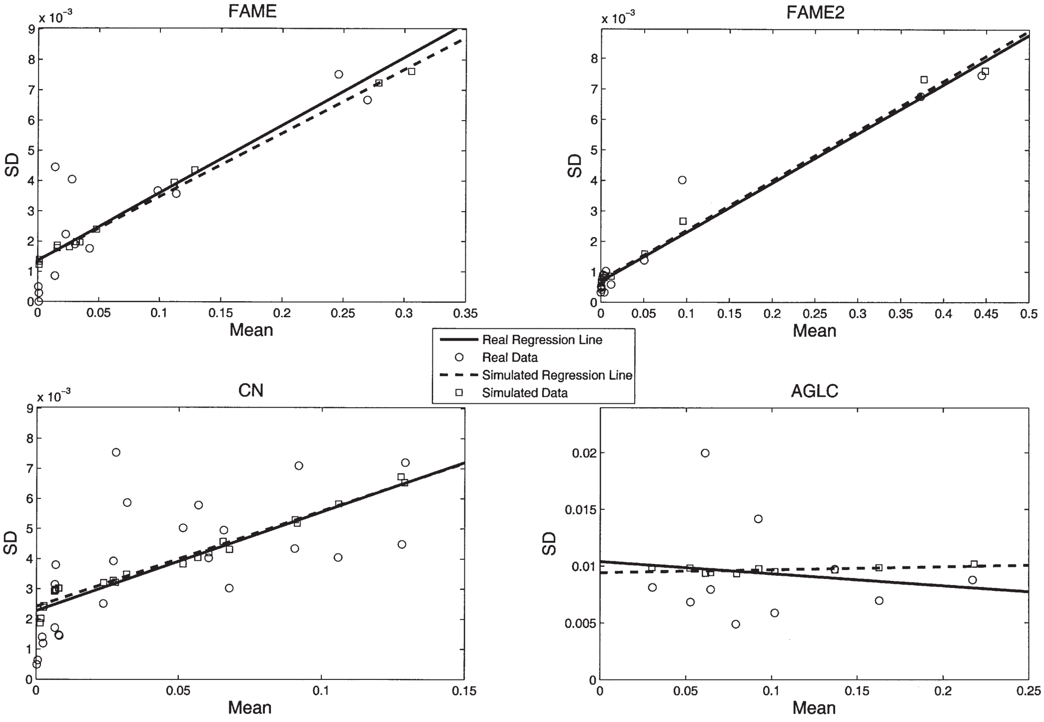

FIG. 5. Scatterplots of the measured (circles and solid lines) and simulated (squares and dashed lines) means and SD of the FAME (upper left), FAME2-pos (upper right), carbon number (CN) (lower left) and AgLC (lower right). The measured data were obtained by measuring a single sample several times (over a time period of 1 yr): FAME: 24; FAME2-pos: 18; CN: 40; and AgLC: 26. Each figure also shows the regression lines corresponding to the data. The noise parameters determined are listed in Table 1. The simulated data represent the results of 1,000 simulated $\mathbf{y}_n$ vectors.

relatively flat TP score around the range at which the maximum was scored. Furthermore, one can see that the EM method scored fewer TP over the complete range of threshold levels. Most interesting are the maximal numbers of TP scored, which are listed in Table 2. This table lists the number of times the methods selected the correct combination of interesterified RM and the maximum number of TP scored for either one of the qualitative strategies. Table 2 shows that the number of times the correct combination of RM was selected was, at the most, 497 for the L1 method and, at the least, 416 for the EM method. This is an unexpected result since L3 is, in principle, better equipped to estimate in(RM), as it allows negative fractions for them. The explanation lies in the fact that the four in(RM) fractions can move freely through the $\mathbf{w}_e(\alpha)$ space. The L3 method is the only one that can theoretically find any point on the trajectory, but none of the methods actually restrict the fractions to a point on the four (scaled) curves (Fig. 3). For the brute force search with the L3 method, we observed that the difference between the lowest and the next-lowest ε_y was smaller than this difference for the L1 and L2 methods. This was caused by the flexibility of the L3 method and more often resulted in an overfitting of the data in a noisy situation.

Overall, the maximum numbers of TP scored for the Niekerk and Vliet methods were close. Since the results were not significantly in favor of either method, we chose to use the Vliet method in the remainder of the paper.

When a method scores the same number of TP for different threshold levels, it can be favorable to choose the lowest threshold level. This is because a lower threshold level will result in more cases in which a surplus RM is estimated to be present, which might be preferable to a higher chance that RM are left out erroneously. The choice in this trade-off will depend on the frequency and cost associated with these two cases of incorrect qualitative estimation.

Quantitative FBR. The four methods were compared quantitatively in terms of ε_w as defined by Equation 25. The comparison is based on 10,000 random $\mathbf{w}_r$. Three different simulation regimes were used: The estimations were made first using 10,000 $\mathbf{y}_n$, second by independently generating two $\mathbf{y}_n$ for each $\mathbf{w}_r$ (thus, 20,000 $\mathbf{y}_n$ in total) and using the average of the two as input for the estimation, and third by using the previous procedure but with four

TABLE 1
List of the Determined Noise Coefficients[a] for All Four Analytical Methods

Method	c_1	c_2
FAME	0.0320	0.0015
FAME2-pos	0.0250	0.0008
CN	0.0500	0.0030
AgLC	0	0.0100

[a] The coefficients refer to Equation 10. CN, carbon number.

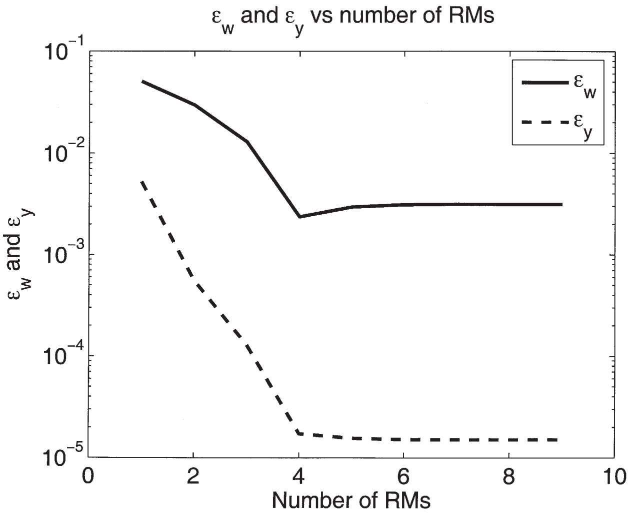

FIG. 6. Plot showing $\overline{\varepsilon_w}$ (solid line)as well as $\overline{\varepsilon_y}$ (dashed line) as a function of the number of RM. The plot was obtained by applying the backward elimination procedure for 2,000 random $\mathbf{w}_r$ with four RM. The backward elimination was started with nine RM down to one for L3. For abbreviation see Figure 2.

independently generated $\mathbf{y}_n$. Using the average of multiple independently generated $\mathbf{y}_n$ as an input for the estimation in fact simulated a noise reduction.

Table 3 shows the best methods in terms of ε_y and $\varepsilon_{w,c}$ based on these 10,000 $\mathbf{y}_n$. One can see that the L3 method was by far superior in terms of ε_y, scoring best in 9,137, 9,382, and 9,535 cases out of the total 10,000 for each of the three different simulation regimes. The method L3 did not perform best for all cases because the qualitative estimations of the four methods can result in different sets of RM. Table 3 also shows the number of TP each method scored for the 10,000 estimations. In addition, it presents a histogram of the TP in terms of the maximum absolute error compared with $\mathbf{w}_r$.

Table 3 shows that the biggest qualitative error is introduced by selecting the wrong combination of RM in the interesterification. Investigation of the incorrect estimates revealed that the most common error made was the swapping of BO and SF. Especially at low values for β and 1 − α (the level of oil present was equal to the product of those two), this occurred frequently. The next main cause of error was the selection of five RM (the correct set of four plus an additional one).

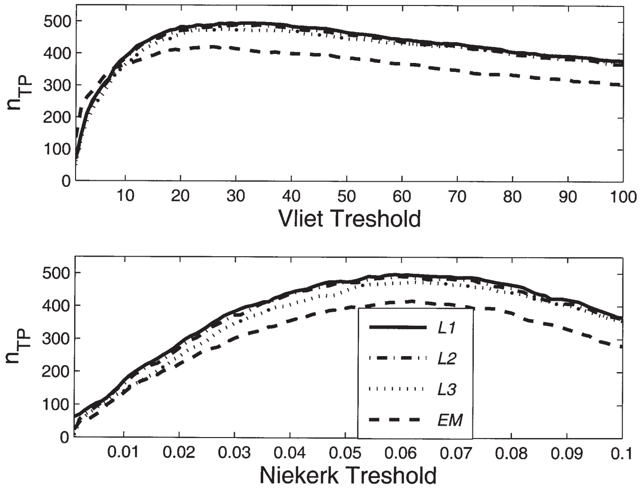

FIG. 7. Plot of the number of true positive (TP) elements (i.e., the number of times the qualitative estimation resulted in the correct set of RM) as a function of the Niekerk threshold (lower figure) as well as the Vliet threshold (upper figure). The Vliet threshold ranged from a 0% increase up to a 100% increase in mean squared error. The Niekerk method had a threshold range from 0 up to 0.10. The results were obtained by simulating the qualitative estimation for 1,000 random $\mathbf{w}_r$.

Table 3 shows that under the noise parameters determined, the L3 method outperformed the others based on ε_w. Using a single $\mathbf{y}_n$, the L3 method found a qualitatively correct solution with a maximum absolute error of <0.01 in 19.0% of the TP cases, which was slightly better than L1 (16.8%) and L2 (18.2%). The EM method, on the other hand, scored considerably worse: In only 9.9% of the cases was the error this small.

The performance increased significantly for all methods when the average of two or four $\mathbf{y}_n$ was used. The performance increase was strongest for the L3 method. Therefore, we propose that the L3 method be used for FBR.

Based on the simulated results, we propose a three-step approach as the current best strategy: Step 1: Determine the pair of RM in the interesterification by a brute-force search. Step 2: Apply backward elimination with the Vliet threshold at 27%. Step 3: Perform a quantitative estimation with L3, preferably using an average of multiple replicated measurements.

Real data. The four methods were compared based on their ability to correctly identify three mixtures. These mixtures were prepared, and FAME, FAME2-pos, CN, and AgLC analy-

TABLE 2
Table Listing the Maxima Found in Figure 7[a]

Method	n_{in}	$n_{\text{TP,V}}$	V	$n_{\text{TP,N}}$	Niekerk
L1	647	494	28%	497	0.058
L2	653	489	30%	492	0.058
L3	625	476	27%	474	0.061
EM	575	420	25%	416	0.061

[a] Column n_{in} indicates the number of times a method found the correct combination of raw materials (RM) in the interesterification (in) based on the qualitative estimation of 1,000 random $\mathbf{w}_r$. Column $n_{\text{TP,V}}$ indicates the actual number of true positives (TP) scored using the Vliet procedure and V the threshold value at which this occurred. For the Niekerk strategy, $n_{\text{TP,N}}$ and N list the same. The numbers correspond to Figure 7.

TABLE 3
Table Presenting the Quantitative Performance of the Four Methods[a]

	ε_y	$\varepsilon_{w,c}$	n_{in}	n_{TP}	<.01	<.02	<.05	<.10	<.20	<1.00
One $\mathbf{y}_n$										
L1	76	2,490	7,977	6,595	0.1820	0.4851	0.8688	0.9854	1.0000	1.0000
L2	713	1,760	7,987	6,484	0.1675	0.4661	0.8590	0.9861	1.0000	1.0000
L3	9,137	3,371	7,698	6,217	0.1900	0.5018	0.8753	0.9883	1.0000	1.0000
EM	74	2,379	7,496	6,010	0.0993	0.3035	0.6865	0.8739	0.9815	1.0000
Total	10,000	10,000								
Two $\mathbf{y}_n$										
L1	43	2,647	8,673	7,735	0.2794	0.6081	0.9325	0.9961	1.0000	1.0000
L2	546	1,726	8,681	7,647	0.2555	0.5868	0.9266	0.9970	1.0000	1.0000
L3	9,382	3,795	8,486	7,303	0.2975	0.6289	0.9366	0.9978	1.0000	1.0000
EM	29	1,832	8,069	6,933	0.1255	0.3433	0.6997	0.8696	0.9864	1.0000
Total	10,000	10,000								
Four $\mathbf{y}_n$										
L1	40	2,879	9,157	8,502	0.3907	0.7098	0.9672	0.9994	1.0000	1.0000
L2	412	1,615	9,161	8,396	0.3486	0.6837	0.9664	0.9996	1.0000	1.0000
L3	9,535	4,313	9,079	8,100	0.4452	0.7553	0.9762	0.9998	1.0000	1.0000
EM	13	1,193	8,371	7,346	0.1465	0.3625	0.7062	0.8618	0.9912	1.0000
Total	10,000	10,000								

[a] The results are listed using one, two, and four $\mathbf{y}_n$. For each method, the number of times it provided the best ε_y and $\varepsilon_{w,c}$ is listed, along with the number of correct interesterification selections (n_{in}) and the number of TP (n_{TP}). For the TP estimates, a tabulated histogram is presented of the maximum absolute error compared with the real composition ($\mathbf{w}_{r,c}$). Results were obtained by 10,000 simulated estimations using the noise model derived and the qualitative estimation parameters as derived in the *Qualitative FBR* section.

ses were performed. The RM themselves were not analyzed. Therefore, the $\mathbf{y}_n$ of the RM in $\mathbf{M}_y$ were not exactly the same. This provided a test bench with analytical noise as well as natural variance noise.

A larger set of RM was used consisting of three oils, i.e., BO, SF, and RP; four MMF, i.e., coconut oil (CN), CN31, PK, and hardened palm kernel oil (PK39); and five HMF, i.e., PO, dfPOs52mp, hardened palm oil (PO58), mfPOs, and hardened rapeseed oil (RP68). Furthermore, all interesterifications based on a mixture of an arbitrary HMF, F_{HMF}, and an arbitrary MMF, F_{MMF}, were considered (thus, 20 possible combinations). Table 4 lists the results, including ε_y and ε_w.

The estimates of Mix 1 were very good: All four methods yielded a qualitatively correct estimate. The differences between the methods were only marginal, slightly favoring the L3 method, as this scored the lowest ε_y and ε_w.

The four estimates of Mix 2 all yielded a correct qualitative estimation. All four estimates were quite close to the composition of the original mixture. However, L1 slightly outperformed the others as it was slightly closer in terms of α and had the lowest ε_w.

TABLE 4
The Mixtures Prepared and Their Estimated Composition Obtained by Each of the Four Methods[a]

Mix 1	30RP/70in(60PO/40PK)	ε_y (10^{-3})	ε_w
L1	26.18SF/73.82in(63PO/37PK)	0.1229	0.0007
L2	26.40SF/73.60in(64PO/36PK)	0.1225	0.0012
L3	26.43SF/73.57in(63PO/37PK)	0.1223	0.0007
EM	26.27SF/73.73in(64PO/36PK)	0.1288	0.0027
Mix 2	50RP/30SF/15in(40PO58/60PK39)/5PO	ε_y (10^{-4})	ε_w
L1	43.59RP/34.14SF/7.38PO/14.89in(45PO58/55PK39)	0.4824	0.0008
L2	44.26RP/34.45SF/6.74PO/14.55in(50PO58/50PK39)	0.4535	0.0019
L3	44.26RP/34.45SF/6.74PO/14.55in(50PO58/50PK39)	0.4535	0.0019
EM	44.22RP/36.16SF/3.89PO/15.72in(54PO58/46PK39)	0.6453	0.0034
Mix 3	60SF/30in(65dfPOs52mp/35PK39)/10PK	ε_y (10^{-3})	ε_w
L1	44.18SF/16.02RP/39.81in(54dfPOs52mp/46PK)	0.6039	0.0165
L2	44.83SF/15.61RP/4.02dfPOs52mp/35.54in(45dfPOs52mp/55PK)	0.5499	0.0205
L3	44.98SF/14.92RP/4.36PK/2.39PO/33.36in(57dfPOs52mp/43PK)	0.5210	0.0128
EM	48.56SF/11.72RP/39.72in(55dfPOs52mp/45PK)	0.8384	0.0150

[a] The last columns indicate ε_y and ε_w. RP, rapeseed oil; in, interesterification; PO, palm oil; PK, palm kernel oil; SF, sunflower oil; PO58, hardened palm oil; PK39, hardened palm kernel oil; dfPOs52mp, palm stearin.

The qualitative results for Mix 3 were worse than for the other two mixtures. Once again, RP was estimated to be present and less SF; combined, they represented the total amount of SF originally present. All four methods estimated the wrong combination of RM in the interesterification: They estimated PK instead of PK39. The L3 method was the only one that indicated the presence of some PK, and incorrectly a small amount of PO as well. The L2 method also found a small amount of dfPOs52mp instead of the PK. Furthermore, the L3 method provided the best estimate for α and scored the lowest ε_w. Therefore, we feel that this is the preferred method.

The results obtained by estimating these mixtures were very promising. More research is needed to draw conclusions on the performance of the four methods in general. It should be noted that these real-life examples suffered from the presence of natural variance: The profiles used in $\mathbf{M}_y$ were not from the same batch as those from which the mixtures were prepared. Furthermore, mixtures containing less than four RM were prepared so as to offer insight into the worst-case performance.

ACKNOWLEDGMENTS

We would like to thank Sijmen de Jong, Hans-Gerd Janssen, Eckhard Flöter, Kees Roos, and Richard Heusdens for their discussions of and suggestions for the work presented.

REFERENCES

1. Ulberth, F., and M. Buchgraber, Authenticity of Fats and Oils, *Eur. J. Lipid Sci. Technol. 102*:687–694 (2000).
2. Precht, D., Detection of Foreign Fat in Milk. II. Quantitative Evaluation of Foreign Fat Mixtures, *Z. Lebensm. Unters. Forsch. 194*:107–114 (1992).
3. Ulberth, F., Detection of Milk Fat Adulteration by Linear Discriminant Analysis of Fatty Acid Data, *J. AOAC Int. 77*:1326–1334 (1994).
4. Ulberth, F., Quantitation of Foreign Fat Milkfat Mixtures by Multivariate Regression of Fatty Acid Data, *J. Agric. Food Chem. 43*:1556–1560 (1995).
5. Simoneau, C., Detection and Quantification of Cocoa Butter Equivalents in Chocolate Model Systems: Analysis of Triglyceride Profiles by High Resolution GC, *Food Chem. 65*:111–116 (1999).
6. Dionisi, F., P.A. Golay, B. Hug, M. Baumgartner, P. Callier, and F. Destaillats, Triacylglycerol Analysis for the Quantification of Cocoa Butter Equivalents (CBE) in Chocolate: Feasibility Study and Validation, *J. Agric. Food Chem. 52*:1835–1841 (2004).
7. Van Niekerk, P.J., and A.E.C. Burger, The Estimation of the Composition of Edible Oil Mixtures, *J. Am. Oil Chem. Soc. 62*:531–538 (1985).
8. van Niekerk, P.J., and R.A. Hasty, Optimisation of a Technique for the Estimation of the Composition of Edible Oil Blends, *Anal. Chim. Acta 223*:237–246 (1989).
9. Abuhadeed, A.M., and A.R. Kotb, A Method for the Quantitative Determination of Individual Oils in a Blend, *J. Am. Oil Chem. Soc. 65*:1922–1926 (1988).
10. De Jong, S., and T.J.R. De Jong, Computer Assisted Fat Blend Recognition Using Regression Analysis and Mathematical Programming, *Eur. J. Lipid Sci. Technol. 93*:532–536 (1991).
11. Boghra, V.R., and V.N. Borkhatriya, Detection of Vegetable Oils in Milk and Milk Fat by a Rapid Method, *J. Food Sci. Technol. (Mysore) 41*:461–464 (2004).
12. Lankmayr, E., J. Mocak, K. Serdt, B. Balla, T. Wenzl, D. Bandoniene, M. Gfrerer, and S. Wagner, Chemometrical Classification of Pumpkin Seed Oils Using UV-Vis, NIR and FTIR Spectra, *J. Biochem. Biophys. Methods 61*:95–106 (2004).
13. Timms, R.E., *Confectionery Fats Handbook: Properties Production and Application*, The Oily Press, Bridgwater, England, 2003.
14. van Vliet, M.H., and G.M.P. van Kempen, Computational Estimation of the Triacylglycerol Composition of Vegetable Fats from Gas and Liquid Chromatography Data, *Eur. J. Lipid Sci. Technol. 106*:697–706 (2004).
15. Buchgraber, M., F. Ulberth, H. Emons, and E. Anklam, Triacylglycerol Profiling by Using Chromatographic Techniques, *Eur. J. Lipid Sci. Technol. 106*:621–648 (2004).
16. Andrikopolous, N., Triglyceride Species Compositions of Common Edible Oils and Methods Used for Their Identification and Quantification, *Food Rev. Int. 18*:71–102 (2002).
17. Laakso, P., Analysis of Triacylglycerols—Approaching the Molecular Composition of Natural Mixtures, *Ibid. 12*:199–250 (1996).
18. Beale, E.M.L., *Introduction to Optimization*, John Wiley & Sons, Essex, Great Britain, 1988.
19. Venkataraman, P., *Applied Optimization with MATLAB Programming*, John Wiley & Sons, New York, 2002.
20. Lawson, C.L., and R.J. Hanson, *Solving Least Squares Problems*, Prentice-Hall, Englewood Cliffs, NJ, 1974.
21. Csiszár, I., Why Least Squares and Maximum Entropy? An Axiomatic Approach to Inference Problems, *Ann. Stat. 19*:2032–2066 (1991).
22. van Kempen, G.M.P., Image Restoration in Fluorescence Microscopy, Ph.D. Thesis, Delft University of Technology, Delft, The Netherlands, 1999.
23. Dempster, A.P., N.M. Laird, and D.B. Rubin, Maximum Likelihood from Incomplete Data *via* the EM Algorithm, *J. Royal Stat. Soc. B 39*:1–33 (1977).
24. Golub, G.H., and C.F. van Loan, *Matrix Computations*, 3rd edn., Johns Hopkins University Press, Baltimore, 1996.
25. The MathWorks, *Matlab 7 Release 14: Reference Manual 2004*, The MathWorks, Natick, MA, 2004, http://www.MathWorks.com.
26. Institute for Reference Materials and Measurements (IRMM), Bcr-164 Reference Sample, Certificate of Analysis, IRMM, Geel, Belgium, 2004.
27. Paquot, C., and A. Hautfenne, *IUPAC Standard Methods for Analysis of Oils, Fats and Derivatives*, 7th edn., Blackwell Science, Oxford, United Kingdom, 1987.
28. Jeffry, B.S.J., Silver Complexation Liquid Chromatography for Fast, High Resolution Separations of Tri-acylglycerides, *J. Am. Oil Chem. Soc. 68*:289–293 (1991).

[Received February 14, 2005; accepted August 15, 2005]