# Mid-Infrared Spectroscopy and Chemometrics for the Authentication of Meat Products

Osama Al-Jowder,[†] Marianne Defernez,[‡] E. Katherine Kemsley,*[,‡] and Reginald H. Wilson[‡]

Institute of Food Research, Norwich Research Park, Colney, Norwich NR4 7UA, United Kingdom, and College of Science, Chemistry Department, University of Bahrain, Isa Town, Bahrain

Mid-infrared (MIR) spectroscopy is used to address certain issues connected with the authentication of beef and ox kidney and liver: is it possible to distinguish muscle from offal tissue; does the condition, cut of meat, or type of offal influence the distinction; can pure minced beef be distinguished from that adulterated with offal? Using partial least squares (PLS) and canonical variate analysis, predictive models are developed to identify MIR spectra of beef, kidney, and liver. Using modified SIMCA, the pure beef specimens are modeled as a single class; this model identifies spectra of unadulterated beef as such, with an acceptable error rate, while rejecting spectra of specimens containing 10−100% w/w kidney or liver. Finally, PLS regressions are performed to quantify the amount of added offal. The prediction errors obtained (±4.8 and ±4.0% w/w, respectively, for the kidney and liver calibrations) are commensurate with the detection limits suggested by the SIMCA analysis.

**Keywords:** *Infrared; spectroscopy; beef; liver; kidney; offal; adulteration*

## INTRODUCTION

Food fraud is not a new problem. Instances of adulteration, that is, the partial substitution of high-value raw materials with cheaper alternatives, have been recorded since ancient times. Some infamous cases of adulteration have led to appalling health problems, such as the substitution of lead oxide for certain spices and, more recently, the addition of antifreeze to wine and of a toxic contaminant to olive oil. However, the outcome of most food fraud raises economic rather than health concerns; no consumer wants to be cheated or to pay more for a commodity than it is worth.

In the case of meat products, fraud may take the form of adulteration with less costly cuts from the same or different animal species, offal, water, or cheaper proteins of animal or vegetable origin. In earlier work (Al-Jowder et al., 1997), we conducted a feasibility study on the use of mid-infrared (MIR) spectroscopy as a rapid method for species identification and quality control of turkey, chicken, and pork mince. Infrared spectroscopy represents an attractive option for quality screening, because it is rapid, low-cost, and noninvasive. The MIR region in particular provides information on a very large number of analytes, and the absorption bands are sensitive to the physical and chemical states of individual constituents. The high spectral signal-to-noise ratio obtained from modern instruments means that even constituents present in quite low concentrations can be detected, as well as subtle compositional differences between multiconstituent specimens.

The work reported in the present paper examines the use of MIR spectroscopy for detecting the adulteration of raw, minced beef with certain types of offal obtained from the same species, specifically, ox kidney and liver. The following issues are addressed: using MIR spectroscopy, is it possible to distinguish beef muscle from offal tissue types; if so, then is the distinction influenced by the cut of the beef, the nature of the offal, or the specimen condition (fresh or frozen−thawed); and finally, is it possible to distinguish pure minced beef from that adulterated with some quantity of offal and with what limit of detection?

## EXPERIMENTAL PROCEDURES

**Specimens.** Specimens of beef, ox kidney, and ox liver from different animals were purchased from local food retailers. The beef muscle tissue included specimens of three different cuts: brisket, neck, and silverside. All specimens were minced on the day of purchase, using a Krups coffee blender, which was carefully washed and dried between each preparation, using detergent, 0.2% Triton-X solution, and distilled water. All minced specimens were divided into two unequal-sized portions. The smaller of these was designated for use as a "fresh, unadulterated" specimen and stored in the refrigerator before spectroscopic acquisition. The larger portion was stored at −30 °C, for later spectroscopic analysis as a "frozen−thawed, unadulterated" specimen and for preparation of mixtures of beef and offal (the "adulterated" specimens). The mixtures were prepared by combining an amount in the range of 10−90% w/w of either kidney or liver with randomly selected specimens of each of the three cuts of beef.

**Instrumentation and Spectral Acquisition.** All spectra were collected on a Spectra-Tech (Applied Systems Inc.) Monitir Fourier transform infrared (FTIR) spectrometer system, fitted with a sealed and desiccated interferometer and a room temperature deuterated triglycine sulfate (DTGS) detector. An attenuated total reflectance (ATR) accessory was built into one of two dedicated sampling stations. The accessory comprised transfer optics within a desiccated chamber, sealed from the atmosphere by two potassium bromide windows. Through these windows, the infrared radiation was directed into a detachable ATR element, which could be removed for cleaning without ingress of water vapor into the spectrometer.

* Author to whom correspondence should be addressed (fax +44 1603 507723; e-mail kate.kemsley@bbsrc.ac.uk).
[†] University of Bahrain.
[‡] Institute of Food Research.

10.1021/jf981196d CCC: $18.00 
Published on Web 07/24/1999

**Table 1. Summary of the Specimen Types and Spectra Recorded**

| specimen type | no. of specimens | spectra recorded | |
|---|---|---|---|
| | | fresh | frozen−thawed |
| beef (brisket) | 19 | 38 | 38 |
| beef (neck) | 21 | 42 | 42 |
| beef (silverside) | 20 | 40 | 40 |
| ox kidney | 21 | 42 | 42 |
| ox liver | 22 | 44 | 44 |
| total | 103 | 206 | 206 |
| mixtures of beef (brisket, neck, or silverside) with 10, 12, 14, ..., 90% w/w ox kidney | 41 | 0 | 41 |
| mixtures of beef (brisket, neck or silverside) with 10, 12, 14, ..., 90% w/w ox liver | 41 | 0 | 41 |
| total | 82 | 0 | 82 |

**Table 2. Allocations of the Spectra to Training and Test Sets for the PLS/CVA, SIMCA, and PLS Regression Work**

| | training set | test set |
|---|---|---|
| PLS/CVA Analysis a | | |
| beef (all cuts) | 160 (40)[a] | 80 (20) |
| kidney | 56 (14) | 28 (7) |
| liver | 56 (14) | 32 (8) |
| PLS/CVA Analysis b | | |
| brisket | 52 (13) | 24 (6) |
| neck | 56 (14) | 28 (7) |
| silverside | 52 (13) | 28 (7) |
| SIMCA Analysis | | |
| beef (all cuts) | 160 (40) | 80 (20) |
| beef adulterated with kidney | 0 | 41 (41) |
| kidney | 0 | 84 (21) |
| beef adulterated with liver | 0 | 41 (41) |
| liver | 0 | 88 (22) |

| PLS Regressions | internal cross-validation |
|---|---|
| beef adulterated with kidney | 41 (41) |
| beef adulterated with liver | 41 (41) |

[a] Numbers in parentheses indicate the number of independent specimens giving rise to these spectra.

The ATR element was a (nominal) 11-reflection zinc selenide crystal mounted into a plate with a shallow trough for specimen containment. The crystal geometry was a 45° parallelogram with mirrored angle faces.

Fresh and frozen−thawed unadulterated specimens were each further divided into two and infrared spectra recorded of both portions. A single infrared spectrum only was recorded of each mixture. A summary of the specimen types and spectra recorded is given in Table 1. For spectral acquisition, the specimens were spread directly onto the ATR element. The ATR plate was thoroughly cleaned between acquisitions by removing the specimen with tissue and cleaning with 0.2% Triton X-100 solution and distilled water. This procedure was found to efficiently remove traces of fat from the ATR crystal. Spectra were recorded from 800 to 4000 $cm^{-1}$, at a nominal resolution of 4 $cm^{-1}$. For each spectrum, 64 interferograms were co-added and a triangular apodization was employed before Fourier transformation. The single-beam spectrum of each specimen was converted into absorbance units using a single-beam background of the clean ATR plate collected under identical conditions. All absorbance spectra were truncated to 470 data points in the range of 900−1800 $cm^{-1}$ (the "fingerprint" region; this is often the most useful part of the MIR spectrum). Baseline correction was carried out using a single definition point at 1800 $cm^{-1}$.

**Chemometric Analysis.** Four distinct investigations were undertaken. For the first three, data analysis was carried out using the Win-DAS (Wiley, Chichester, U.K.) software package. For the fourth investigation, data analysis was carried out using the MatLab matrix programming language (The Mathworks Inc., Natick, MA).

*Exploratory Analysis.* In addition to visual examination of the spectra, a well-known "data compression" technique was employed, called principal component analysis (PCA) (Jolliffe, 1986). A primary aim of PCA is to transform data comprising measurements of many properties or "variates" (here, these are measurements of spectral absorbance at several hundred infrared frequencies) into a new data set of much more manageable size. The transformed variates are known as principal component (PC) "scores" and are ordered such that the first few contain most of the information that was spread across all of the original data. This reduction in complexity is helpful for exploration of large and complex data sets. Patterns or clusters may be revealed that were not apparent from viewing the original data, although, crucially, no predefined structure is imposed; in other words, PCA is not a modeling technique.

*Partial Least-Squares (PLS)/Canonical Variate Analysis (CVA) Modeling.* Following data exploration, a combination of chemometric methods was used to model two potential group structures within the data set. The first model sought to distinguish between the beef muscle, kidney, and liver specimen types. The second model involved the spectra of muscle tissue only and sought to differentiate the three different cuts of beef. The chemometric approach used for this modeling work comprised two further data compression methods: PLS (Martens and Naes, 1989a) and CVA (Krzanowski, 1988). PLS is best-known for its use in calibration applications, in which the relationship between two data sets (spectroscopic and concentration data, for example) is modeled using a series of local least-squares fits; it is used in this way in the fourth of the data analyses, discussed below. However, PLS can also be used in discrimination work, in which case the second data set comprises one or more "dummy" variates, which represent the proposed group structure. This approach was taken in the present analysis. The scores produced in PLS also reduce complexity in the data set and can be passed as variates to subsequent procedures.

CVA also makes use of information on the proposed group structure; CV scores by definition exhibit maximum separation and minimum spread of the predefined groups. CVA is also a data compression technique, insofar as it rearranges the information in a data set so that usually only the first few scores are of interest. However, it cannot be applied directly to "high-dimensional" data (in which the number of variates exceeds the number of spectra), and for this reason, PLS is used as a precursor to CVA in the present work; subsets of the available PLS scores only are passed to the CVA procedure. Because in both cases, there are only three proposed groups, only two CV dimensions exist, and the characteristics of the CVA transformation mean that a plot of the CV scores provides the best possible representation for highlighting any differences between the groups. CVA can also be used to make definitive statements about the group membership of individual spectra, through the use of tolerance regions constructed in the CVA coordinate system.

Like all modeling methods, PLS and CVA are best carried out in "training" and "test" phases. In the training phase, the CVA transformation is applied to PLS scores obtained from data allocated to a "training set". An impression of the performance of the PLS/CVA model is obtained from the CV score plot and tolerance regions, but it is also important to carry out a test phase, in which the model is applied to a set of independent data not used the model development (the "test set"). The aim of the test phase is to show that the model is neither over- nor underfit and is able to generalize successfully to data other than that used in the training phase. The PLS/CVA modeling work reported here was carried out in training and test phases. The allocations of the available data to training and test sets are given in Table 2. In both analyses,

**Table 3. Typical Compositional Data for Beef, Kidney, and Liver Obtained from the Literature**

| specimen type | water (% w/w) | protein (% w/w) | fat (% w/w) | carbohydrate (% w/w) |
|---|---|---|---|---|
| beef, lean muscle tissue (av values obtained from 6 different cuts) | 74 | 20.3 | 4.6 | 0 |
| beef, fatty tissue (av values obtained from 6 different cuts) | 24 | 8.8 | 66.9 | 0 |
| brisket (av values obtained from 18 specimens, 77% lean) | 62.2 | 16.8 | 20.5 | 0 |
| silverside (av values from 10 specimens, lean) | 72.2 | 23.8 | 4.3 | 0 |
| ox kidney (av values obtained from 18 specimens) | 79.8 | 15.7 | 2.6 | 0 |
| ox liver (av values obtained from 30 specimens) | 68.6 | 21.1 | 7.8 | 2.2 (present as glycogen) |

it was decided to apportion approximately two-thirds of the spectra in each proposed group to the training set and one-third to the test set, ensuring that each family of four spectra (two "fresh", two "frozen−thawed") arising from a single original specimen was apportioned to either the training or test set. Within these constraints, the allocations were made randomly.

*SIMCA Class Model for Spectra of Pure Beef Muscle.* The aim of this analysis was to develop a protocol for detecting the adulteration of minced beef with either of the two offal types. This kind of problem is sometimes called "asymmetric", because it involves in essence two dissimilar groups of specimens: one well-defined category of interest (here, the pure minced beef) and a second group comprising a much broader range of specimens (liver, kidney, or beef mince adulterated with some quantity of these). Asymmetric problems can be tackled with discriminant analysis methods such as PLS/CVA. However, the class modeling approach is conceptually more appealing; it places no constraints on the distributional characteristics of the broad group, whereas in CVA, all groups are ideally normally distributed. In class modeling, the assumption is made that data from the well-defined category constitute a single normally distributed class, from which it follows that certain quantities derived from the data will also be distributed in a particular way. Tests applied to these quantities are then employed to determine whether individuals can be regarded as members of the class or should be rejected as nonmembers.

The class modeling approach used in the present work is based on the SIMCA method, in which PCA plays an integral part. Let a spectrum be represented by a $(1 \times d)$ vector $\mathbf{x}$, where $d$ is the number of variates. In SIMCA, this vector is regarded as a composite of three quantities: the "class center" $\boldsymbol{\mu}$ (which is the mean training set spectrum), a product term $\mathbf{z}_{(r)}\mathbf{P}^{\mathrm{T}}_{(r)}$, obtained by multiplying a $(1 \times r)$ vector $\mathbf{z}_{(r)}$ of PC scores for the spectrum by an $(r \times d)$ matrix $\mathbf{P}_{(r)}$ of PC loadings computed from the training set, and a $(1 \times d)$ vector $\mathbf{e}$ of "residuals" for that spectrum:

$$\mathbf{x} = \boldsymbol{\mu} + \mathbf{z}_{(r)}\mathbf{P}^{\mathrm{T}}_{(r)} + \mathbf{e} \qquad (1.1)$$

The first of the SIMCA class membership tests concerns the subset of scores $\mathbf{z}_{(r)}$. In geometric terms, a "box" is defined in the PCA coordinate system, within which the scores for a spectrum must lie if it is to be accepted as a genuine class member. The boundaries of the box in each PC dimension are given by the critical values at the required probability level in a suitable statistical test, for example, Student's $t$ test.

The second class membership test involves the vector of residuals $\mathbf{e}$. The size of the elements in this vector can be summarized by calculating the sum of their squares, to give the "residual-sum-of-squares" for the spectrum, $\mathrm{RSS_s}$. A similar quantity can be defined that summarizes the residual for the entire training set, $\mathrm{RSS_{class}}$. The ratio $\mathrm{RSS_s/RSS_{class}}$ is in effect a ratio of two variances, suggesting that the well-known $F$ test may provide an indicator of class membership. In practice, however, some researchers have found that the proper critical value in the $F$ test is too restrictive (Defernez, 1996). An empirical criterion of $\mathrm{RSS_s/RSS_{class}} < 3$ has been proposed as an alternative test for class membership (Kemsley, 1998), and this "modified $F$ test" is used in the present work. To be accepted as a genuine class member, a spectrum must pass both the collective $t$ tests and the modified $F$ test.

In class modeling, the training set comprises data from the single class of interest only. However, in the test phase of the analysis, the class model should be challenged with both known members and nonmembers of the class, to examine its ability to respectively accept or reject each type of individual correctly. The allocation of specimens to the training and test sets for the present analysis is given in Table 2.

*Calibration Using PLS Regression for the Offal Content of Muscle/Offal Mixtures.* The purpose of this analysis was to determine whether the amount of added kidney or liver could be quantified in the MIR spectra of the mixture specimens only. For this work, we have used PLS regression. In this method, the spectral data are transformed to a set of new variates (the PLS scores), using the concentration data to determine the parameters of the transformation. A subset of the scores is then used in a conventional multiple least-squares regression onto the concentration data. As in all regression methods, it is important to ensure that the model obtained is neither over- nor underfit by conducting a validation procedure. In the previous analyses, this was achieved through separate training and test phases, but in the present analysis, we will use an alternative approach called "internal cross-validation" (ICV). In ICV, one spectrum is omitted from the data set, and the remaining data are used to obtain a model, which is then applied to the excluded, "test" spectrum. This process is repeated as many times as there are spectra in the data set, omitting each spectrum in turn; this produces a complete series of predictions for the entire data set. ICV is an acceptable alternative to the training/test approach, and it has the advantage of being particularly suitable when small numbers of specimens are involved (Defernez and Kemsley, 1997). In the present work, ICV is carried out for PLS regression models over a range of different dimensionalities (i.e., numbers of PLS scores), and the optimal subset identified by the first minimum in the error of prediction.

## RESULTS AND DISCUSSION

**Exploratory Analysis.** The chemical composition of different beef cuts can vary quite considerably, depending on the proportions of fat and lean tissue present; the same cut taken from different animals can also show variability. The major constituents are water, protein, and fat. Although quantitative determination was not the object of the present study, it is useful to examine typical values for these constituents published in the literature (McCance and Widdowson, 1991; Chan et al., 1995). An indication of the water, protein, and fat contents of lean and fat tissue is given in Table 3. A typical cut contains a mixture of lean and fat and so will have a composition intermediate between these extremes. Typical compositional data for two of the cuts used in the present study, brisket and silverside, are also given in the table. No information could be obtained for neck. Data are also available in the literature for the two types of offal used in the study, kidney and liver. There is an important point of note to be made from examination of these values: in addition to water, protein, and fat, liver also contains an appreciable amount of carbohydrate, present in the form of glycogen.

Some of these compositional characteristics are reflected in the infrared spectra of the different specimen

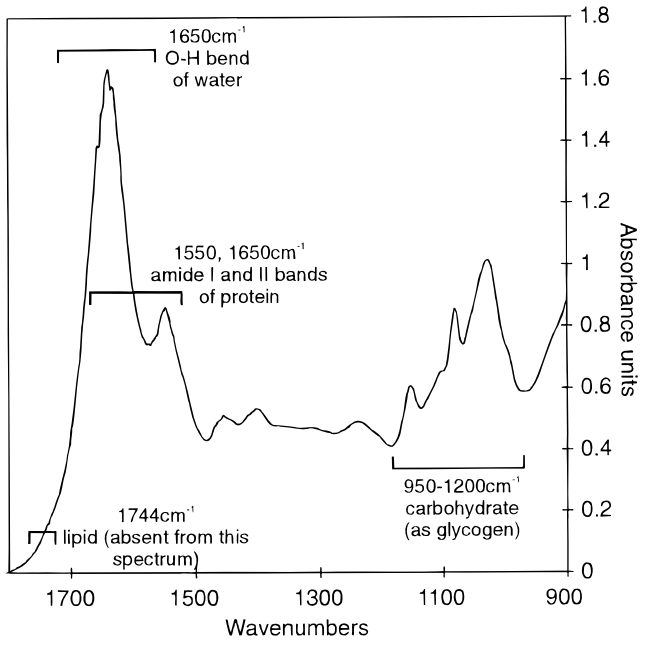


**Figure 1.** Fingerprint region of a spectrum of one selected liver specimen.

types. Figure 1 shows the fingerprint region of a spectrum of one selected fresh liver specimen. Although the spectral features are quite overlapped, it is nevertheless possible to attribute certain prominent features to individual chemical species; bands associated primarily with water, fat, protein, and carbohydrate are indicated. Figure 2 shows the fingerprint region of the complete collection of fresh, unadulterated specimens (the spectra of the frozen−thawed specimens were highly similar and in the interests of clarity are not shown). The most noticeable variation between these sets of spectra is the group of features in the region 1000−1200 $cm^{-1}$ present in the spectra of liver and attributable to the glycogen content of these specimens. Although the other specimen types are not thought to contain carbohydrate, their spectra nevertheless show some features in this region, which are also quite variable, albeit much smaller in size. Whereas carbohydrates exhibit strong absorptions, other chemical species including fats give rise to minor bands in this region, and these small features can be attributed to other absorbing constituents of beef and kidney. The fat content indicated by the slight shoulder at 1744 $cm^{-1}$ appears to be highest in the brisket and silverside specimens, lower for the cuts of neck, and lower still or absent for the offal. This is consistent with what we might expect from the published compositional data. The spectra of the three different cuts of beef are visually very similar. There is generally more variation within the brisket and silverside groups, which may imply greater compositional variability. However, some caution needs to be exercised in interpreting absolute spectral intensities, due to the problems of analyzing semisolids using ATR. Irreproducible contact between the specimen and the ATR crystal is one potential source of unwanted spectral variability. In addition, it can be difficult to use ATR for compositional analysis of specimens that are heterogeneous on the micrometer scale, although this is not of primary concern in the present work.

PCA was applied to the complete collection of spectra of unadulterated specimens, fresh and frozen−thawed. The first five PC scores only were retained for examination to make exploration as simple as possible. The variances of these accounted collectively for >99% of the total variability in the data set (respectively, 87, 8, 2, 1, and 1%), so we can be confident that only a very small proportion of the total information content will be neglected by discarding the remaining scores.

Plots of the first five PCs against one another were examined to look for patterns or clusters in the data that corresponded to the different specimen types. Figure 3a shows a plot of the first versus fifth scores, which exhibited the clearest division of the spectra according to specimen type. The kidney and liver groups in particular are well-separated. The groups corresponding to the three different cuts of beef are much more overlapped, although there is still some indication that the specimen types may be distinguished. However, the fresh and frozen−thawed specimens within any one type did not form separate groups, neither in this score plot nor in any of the other plots examined. Figure 3b shows an expansion of the region indicated in Figure 3a, in which are plotted the scores for the silverside specimens only. We conclude that any influence the condition of the specimens may have is small compared with the other sources of variability in the data set.

However, earlier work (Al-Jowder et al., 1997) had indicated that MIR spectroscopy can readily distinguish between fresh and frozen−thawed turkey, chicken, or pork, by applying PCA to the data from each specimen type separately. This approach was applied to each of the five specimen types in the present study, and plots of the PC scores were examined. Some distinction between the fresh and frozen−thawed specimens could be found, although, in all cases, the discrimination was not as great as had been found in the earlier work on other meat species. Figure 4 presents two typical score plots, showing the PC scores that best highlight the distinction between fresh and frozen−thawed specimens of silverside and kidney.

Returning to Figure 3a, we see that the first PC only provides quite a good discriminator between the beef, kidney, and liver specimens. In such instances, where only one or very few PC scores demonstrate a pattern of interest, it can be worthwhile examining the corresponding PC "loadings". In mathematical terms, each PC score is a weighted sum of the original variates. It is produced by multiplying the original spectra by a vector of "weights": these vectors are usually known as loadings. A large weight (positive *or* negative) on a particular variate indicates that it makes an important contribution to the associated score. Furthermore, weights that are close to zero indicate that a particular variate has little influence on the score. When one is dealing with spectroscopic data, it makes sense to plot the loadings themselves as spectra, that is, as response versus spectral frequency. It is often found that features can be identified in the loadings that were also present in the original spectra, and sometimes these can be ascribed to individual absorbing species. In Figure 5, we present the first five PC loadings from the present data set. The first loading represents primarily protein and carbohydrate. The latter is not surprising, because we have already noted that the spectra of liver can be distinguished visually by the presence of bands arising from glycogen in the region 950−1200 $cm^{-1}$. However, the positive features in the region 1550−1650 $cm^{-1}$ that can be identified with the amide I and II protein

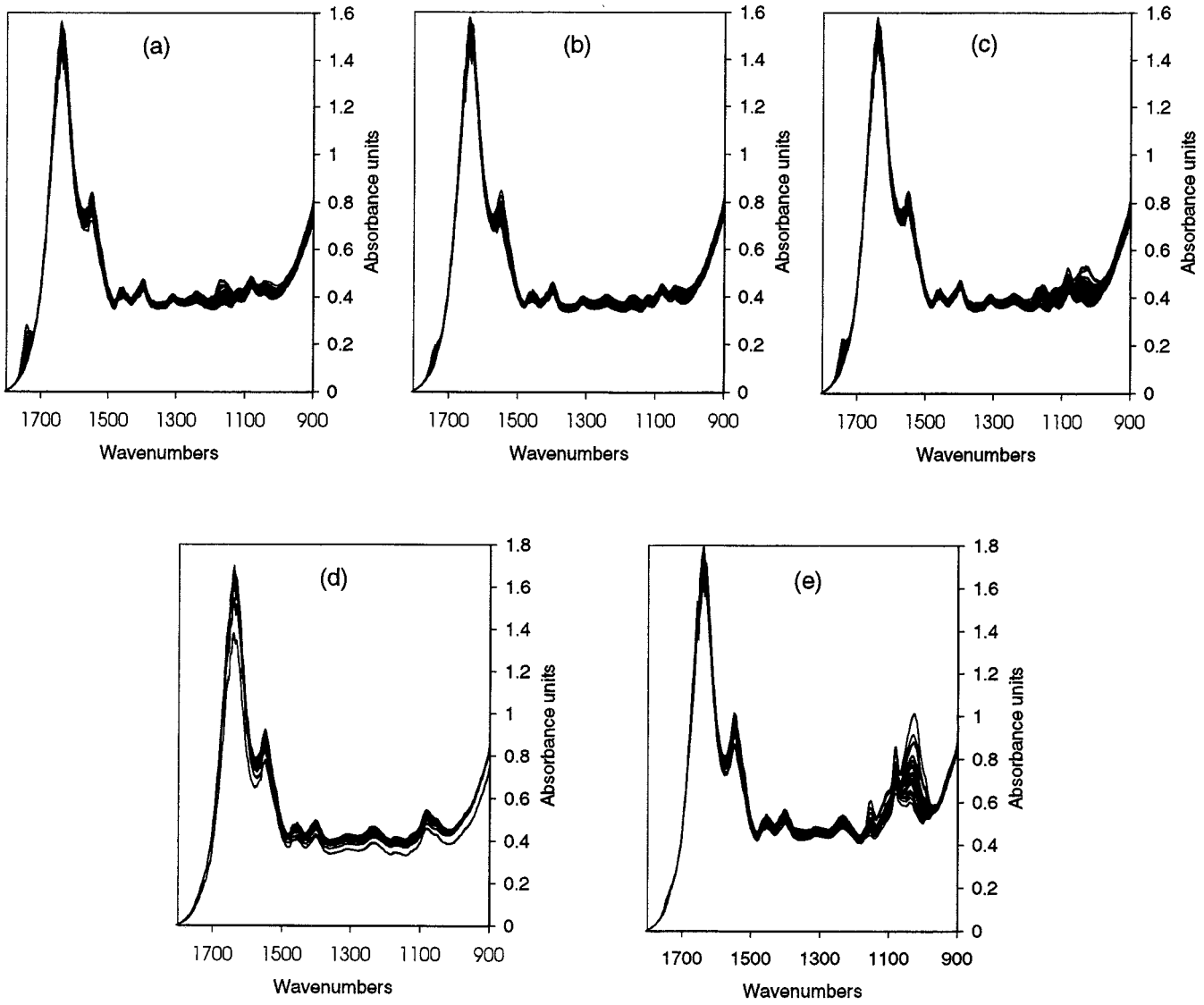


**Figure 2.** Fingerprint region of the complete collection of fresh, unadulterated specimens: (a) brisket; (b) neck; (c) silverside; (d) kidney; (e) liver.

absorptions show that protein content also helps to distinguish the different specimen types.

The exploratory analysis has shown that it is relatively easy to identify the spectra of beef, kidney, and liver. There is also some indication that it may be possible to distinguish between the three different cuts of beef. In the next section, we investigate whether these two potential group structures can be effectively modeled. Specimen condition appears to have a more minor effect on the spectra: to reveal a distinction between the fresh and frozen−thawed groups, the data from each specimen type need to be examined separately. In light of these findings, no further attention is paid to the condition of the specimens in the work that follows, other than to note that spectra of fresh and frozen−thawed individuals are represented equally within all defined training and test sets.

**PLS/CVA Modeling.** The first group structure to be modeled using PLS and CVA comprised three predefined groups: beef (all cuts), kidney, and liver. It was felt that in light of the exploratory analysis, the three different cuts of beef should be placed together in one group for this analysis, because these specimens are clearly much more similar to one another than to the offal types. PLS data compression was applied to the training set as defined in Table 2.

The next step in the analysis was to apply CVA to a subset of the PLS scores. The number of scores to pass as variates to the CVA is an important issue. There are some "rules-of-thumb" that can help to ensure that the models produced are useful and valid. In general, the analysis should be made as simple as possible; variates (that is, PLS scores) should not be included just because they are available. To avoid overfitting, the number of PLS scores used in the CVA should not exceed one-sixth the number of independent specimens in the training set. In the present analysis, CVA was applied to the first three PLS scores only. Figure 6a shows the CV score plot obtained, with the 95% tolerance regions [as described in Kraznowski (1988)] for each group indicated. The three specimen types are well-separated, with no overlap between the 95% tolerance regions for each group. Because the training set comprised data from 68 independent specimens (which produced 272 spectra), we can be reasonably confident that the model is not overfit; however, a test phase will further prove its validity. Figure 6b shows the results obtained from applying this model to the test set. The 95% tolerance

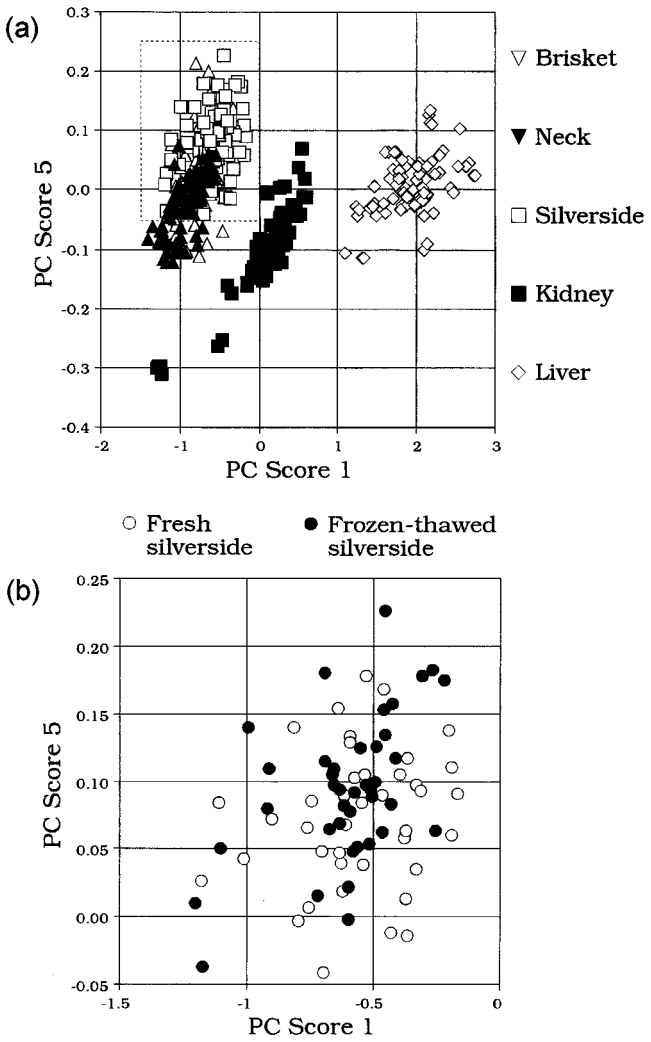


**Figure 3.** Results of the PCA of the entire unadulterated data set: (a) first versus fifth PC scores; (b) expansion of the marked region in (a), showing scores for the silverside specimens only.

regions established in the training phase are also marked on this figure. The pattern of CV scores is highly similar to that obtained for the training set, indicating that the model is not overfit and is able to generalize. We conclude that a successful predictive model has been obtained, and MIR spectroscopy is readily able to differentiate the beef, kidney, and liver specimen types.

The same PLS/CVA approach was applied to the training set comprising data from the three different cuts of beef only. In this case, the first nine PLS scores were required to produce a CV score plot that adequately separated the specimen types (Figure 7a), although the 95% tolerance intervals for the brisket and silverside groups still exhibited some overlap. Furthermore, because the number of independent specimens is now only 40 (160 spectra), there is a strong possibility that this model is somewhat overfit. Indeed, proceeding to the test phase, we find that the performance of the model on the test set is appreciably poorer (Figure 7b). However, there are still some systematic differences; spectra of the neck specimens in particular form a distinct group that can be readily distinguished. Although the brisket and silverside groups exhibit more scatter and overlap, some degree of separation is nevertheless present. These findings are consistent with the observed differences in the original spectral data.

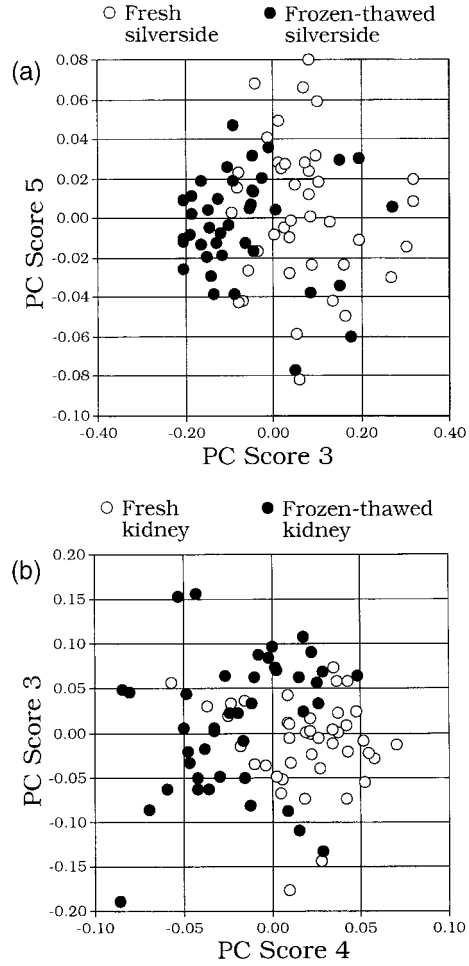


**Figure 4.** Typical score plots obtained from separate PCAs applied to spectra of (a) silverside and (b) kidney.

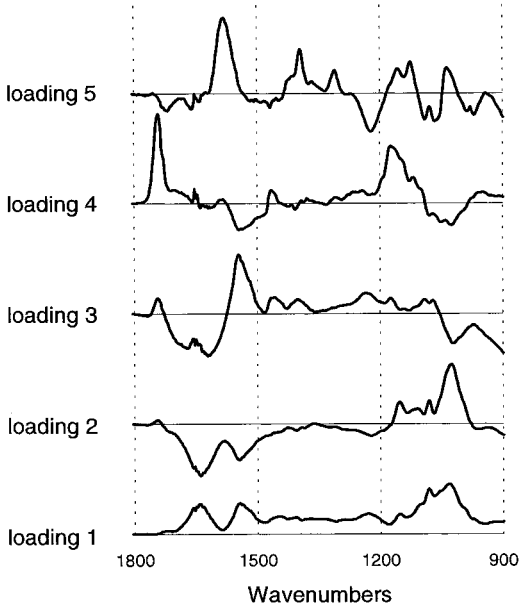


**Figure 5.** First five loadings obtained from the PCA of the entire unadulterated data set. Vertical scale is the same in all cases, but the loadings are shown offset for clarity.

To determine reliably the boundaries of the brisket and silverside groups, the number of specimens analyzed needs to be increased; furthermore, it is possible that some inherent overlap will always be present. We conclude that certain cuts of beef can be identified using

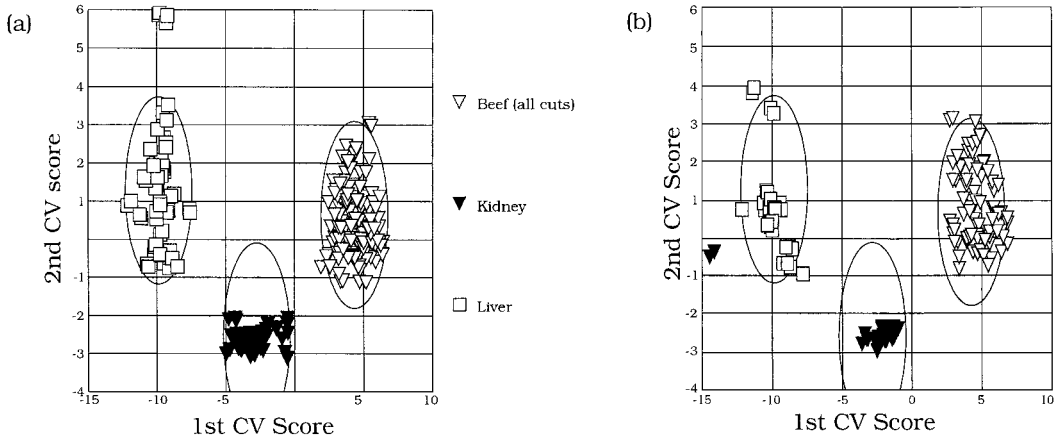


**Figure 6.** CV score plot for the beef, kidney, liver model (using three PLS scores): (a) training set; (b) test set.

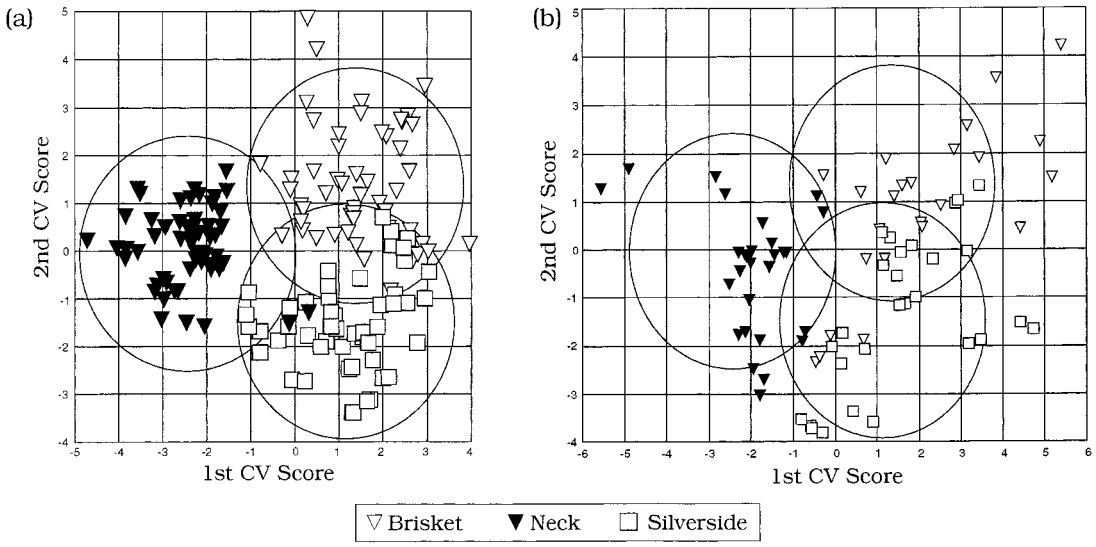


**Figure 7.** CV score plot for the brisket, neck, silverside model (using nine PLS scores): (a) training set; (b) test set.

MIR spectroscopy, although for others the distinction may be less clear. Furthermore, the differences among these specimen types are considerably less than between the beef and offal types.

**SIMCA Class Model for Spectra of Pure Beef Muscle.** The first step in the SIMCA analysis was to apply PCA to the training set comprising data from specimens of pure beef mince only. The variance associated with each PC was examined. The first six PC scores were found to have variances representing ≥1% of the total information content (respectively, 56, 22, 13, 5, 2, and 1%) and collectively accounted for ~99% of the total variance. These thresholds are sometimes used to determine which PCs to retain; inclusion of scores with very small variances is not desirable, because it is likely to make the resultant class models too specific and unable to generalize. We have used these criteria for subset selection in the present work.

A SIMCA class model was created. This comprised two steps. First, a "box" was defined in the first six dimensions of the PCA coordinate system, with the boundaries of the box collectively defined by the *t*-test critical value at the 5% probability level. To be accepted as a genuine class member, scores for an individual spectrum must fall within this box. Second, the quantity $RSS_s/RSS_{class}$ was calculated for each spectrum from the residual PC dimensions. To be accepted as a genuine class member, an individual must also fulfill the criterion $RSS_s/RSS_{class} < 3$. Note that the PCA coordinate system was defined by the training set only, and individuals in the test set were rotated into this coordinate system before application of precisely the same class membership tests.

**Table 4. Results of the SIMCA Class Membership Tests, Expressed as Percentage Acceptance of Each Specimen Type into the "Pure Beef" Class**

| specimen type | accepted by *t* test (%) | accepted by modified *F* test (%) | accepted by both tests (that is, accepted as pure beef) (%) |
|---|---|---|---|
| beef (includes training and test set specimens) | 94 | 95 | 92 |
| kidney | 13 | 0 | 0 |
| liver | 0 | 3 | 0 |

The acceptance rates for each type of specimen are expressed in percentage terms in Table 4. We see that 94% of the spectra of pure beef are accepted by the *t* test as such; this corresponds to a "type I error rate" (incorrect rejections) of 6%, which is commensurate with the probability level in use. A similar acceptance rate of 95% is obtained by the modified *F* test, but it is not possible to comment on whether this outcome is as expected, because the critical value employed is pragmatic rather than probabilistic. With the requirement that an individual passes both class membership tests, we find that 92% of the spectra of pure beef are accepted as such, corresponding to a type I error rate of 8%. In

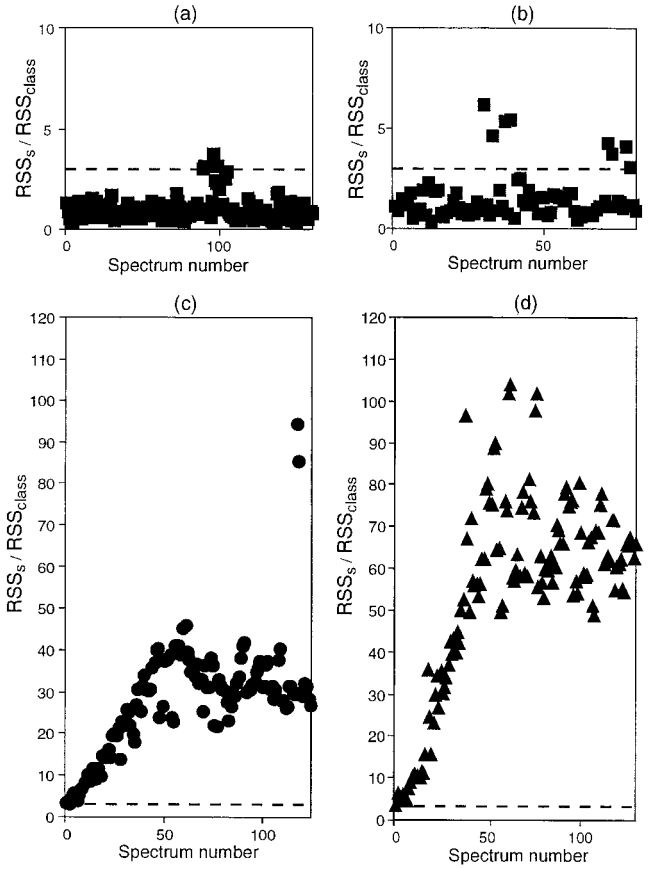


**Figure 8.** Quantity $RSS_s/RSS_{class}$, plotted for (a) training set (■, pure beef specimens); (b) test set (■, pure beef specimens); (c) test set (●, specimens containing 10−100% w/w kidney); (d) test set (▲, specimens containing 10−100% w/w liver). The boundary at $RSS_s/RSS_{class} = 3$ is marked (− − −).

contrast, it is found that none of the spectra of adulterated and "non-beef" specimens are accepted by the SIMCA class model. Therefore, we estimate the detection limits for both kidney and liver in beef to be <10% w/w.

The results of the modified *F* test can be easily represented graphically, and this is done for the complete set of spectra in Figure 8. The quantity $RSS_s/RSS_{class}$ calculated for the training set, which comprises specimens of pure beef only, is shown versus spectrum number in Figure 8a. For ease of examination, the results for the test set are presented separately for each specimen type: Figure 8b shows $RSS_s/RSS_{class}$ for the pure beef specimens, and Figure 8c and Figure 8d show $RSS_s/RSS_{class}$ for specimens containing 10−100% kidney and liver, respectively. Also included on the plots is the threshold value of 3, which marks the accept/reject boundary for the class. It is less easy to represent the SIMCA *t* test in graphical form. For the present analysis, in which six PC dimensions are employed, a two-dimensional score plot can at best depict only a limited projection of the box defining the class. However, this may still provide quite an effective means of visualizing the class model, particularly if certain PCs are especially important for separating the pure and adulterated specimens. In fact, we find that a plot of the first versus second PCs provides the best such representation (Figure 9). Also marked on this figure are the boundaries of the box projected into these two dimensions. From this figure, we can identify 19 spectra of pure beef that are rejected from the class for falling

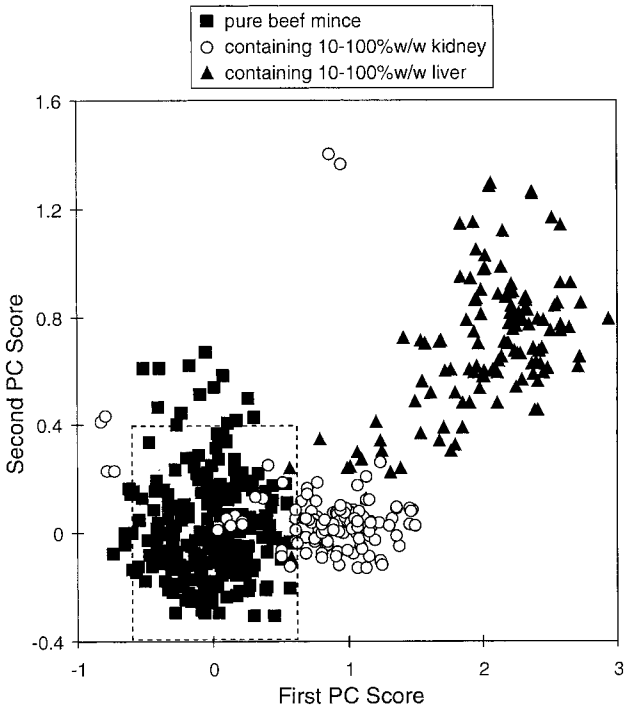


**Figure 9.** Representation of the SIMCA class model in two dimensions of the PC coordinate system.

outside the boundaries of the box. It should be remembered, however, that the spectra of "non-beef" specimens falling within the box boundaries are not incorrectly accepted into the class, because they fail either the *t* test in at least one of the remaining PC dimensions or the modified *F* test.

The "type II error rate" (incorrect acceptances) of 0% is clearly a desirable outcome. However, the relative rates of type I and II errors can be controlled to some extent by the analyst and are therefore somewhat arbitrary. For example, it can be seen from Figure 8 that by increasing the threshold used in the modified *F* test and therefore the number of type II errors, the number of type I errors could be reduced. Similarly, employing a lower probability level in the *t* test could in theory decrease the type I error rate but would probably lead to an increase in the number of type II errors. This tradeoff between the two error types is an issue to be borne in mind whenever class membership tests are defined, and in practice, the acceptance criteria can be adjusted to meet the requirements of the task at hand.

**Calibration Using PLS Regression for the Offal Content of Muscle/Offal Mixtures.** Two separate analyses were carried out, one to calibrate for the quantity of added kidney and the second for the quantity of added liver in minced beef. As a first step, PLS regression was applied to the two data sets comprising 41 spectra in each case (see Table 2). The model dimensionality was varied across a range and the standard error of calibration (SEC) (Martens and Naes, 1989b) recorded. Next, ICV was used to obtain the standard error of prediction (SEP) (Martens and Naes, 1989c) for the same range of model dimensionalities. Both the SEC and the SEP are shown in Figure 10 versus the number of PLS scores used. The optimum model dimensionality was chosen as that producing the first minimum in the SEP. This occurred for the kidney calibration using eight PLS scores and for the liver calibration using four PLS scores. This is consistent with

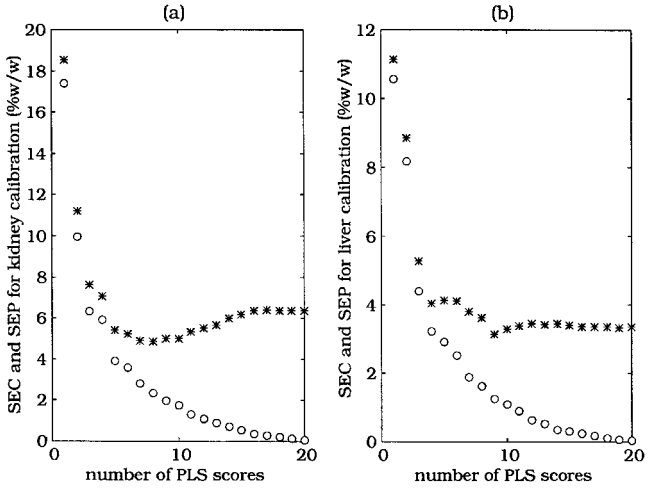


**Figure 10.** SEC (○) and SEP (∗) as a function of the number of PLS scores for (a) kidney and (b) liver added to minced beef.

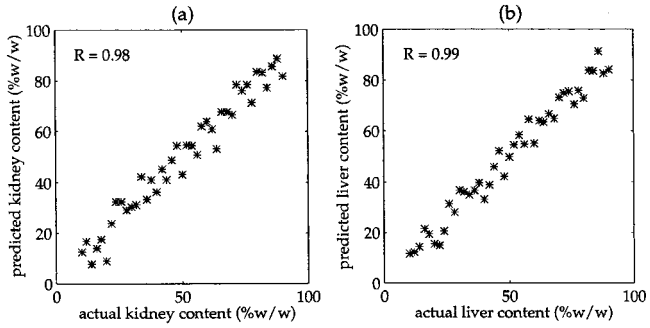


**Figure 11.** ICV predictions for added (a) kidney (using eight PLS scores) and (b) liver (using four PLS scores).

the larger differences observed between the spectra of liver and beef specimen types; in effect, this is an "easier" problem for the regression to tackle. The predictions obtained by ICV are plotted versus the actual concentration of adulterant in Figure 11, for the eight- and four-score models for, respectively, kidney and liver. The correlation coefficients ($R$) between the actual and the predicted values are marked on the plots. The SEP was 4.8% w/w for the kidney calibration and 4.0% w/w for the liver calibration. The definition of the detection limit in regression is somewhat controversial (Miller and Miller, 1993), but the values of the SEP are commensurate with the detection limits suggested by the SIMCA analysis of <10% w/w for both adulterants. Thus, we conclude that PLS calibration enables us to quantify the amount of offal added to the meat cuts, with acceptable precision and accuracy.

## CONCLUSIONS

The work reported here has shown that MIR spectroscopy is useful for a variety of different analyses of minced beef, ox kidney, and ox liver. It is readily able to distinguish between the muscle and offal tissue types. In some instances, the identification can be made directly by visual examination: for example, MIR spectra of liver exhibit bands arising from the small glycogen content, which distinguishes these from the other specimen types. However, chemometric methods are required for the reliable identification of the remaining specimens.

Using the combined technique of PLS/CVA, predictive models were developed to distinguish the beef, kidney, and liver specimens and the three different cuts of beef. It was found that the beef and offal types could be readily distinguished; however, the distinction among the three cuts of beef was much less clear. The neck specimens formed a distinct group, but the silverside and brisket groups could not be entirely separated, and it is possible that there is an inherent overlap between these two specimen types. Exploratory analysis showed that specimen condition (fresh or frozen−thawed) had a relatively minor influence on the spectra, and this factor is not believed to have affected the outcome of the modeling work.

SIMCA class modeling was used to develop a single class model for the pure beef specimens. This model was able to accept spectra of pure beef as genuine class members, with a type I error rate of ∼8%. The SIMCA model was also challenged with beef specimens adulterated with quantities in the range of 10−100% w/w of kidney or liver. All of these specimens were rejected by the model (giving a type II error rate of 0%). This suggests a detection limit for these adulterants of <10% w/w. Finally, PLS regression was used to quantify the amount of added kidney and liver, in two separate calibrations. The values for the SEPs obtained (4.8% w/w for the kidney calibration and 4.0% w/w for the liver calibration) were commensurate with the suggested detection limits obtained by the SIMCA analysis.

## LITERATURE CITED

Al-Jowder, O.; Kemsley, E. K.; Wilson, R. H. Mid-Infrared Spectroscopy and Authenticity Problems in Selected Meats: a Feasibility Study. *Food Chem.* **1997**, *59* (2), 195−201.

Chan, W.; Brown, J.; Lee, S. M.; Buss, D. H. *Meat, Poultry and Game;* Royal Society of Chemistry: London, U.K., 1995.

Defernez, M. Methods based on Principal Component Analysis of Mid-Infrared Spectra: A New Approach for the Classification and Authentication of Fruit Products. Ph.D. Dissertation, The University of East Anglia, 1996; pp 59−77.

Defernez, M.; Kemsley, E. K. The Use and Misuse of Chemometrics for Classification problems. *Trends Anal. Chem.* **1997**, *16* (4), 216−221.

Jolliffe, I. T. In *Principal Component Analysis*; Springer-Verlag: New York, 1986; pp 1−5.

Kemsley, E. K. In *Discriminant Analysis and Class Modelling of Spectroscopic Data*; Wiley: Chichester, U.K., 1998.

Krzanowski, W. J. In *Principles of Multivariate Analysis: A User's Perspective*; Oxford Science Publications, Clarendon Press: Oxford, U.K., 1988; pp 291−306.

Martens, H.; Naes, T. In *Multivariate Calibration*; Wiley: Chichester, U.K., 1989; (a) pp 121−125, (b) p 254, (c) p 251.

McCance, R. A.; Widdowson, E. A. In *The Composition of Foods*; Royal Society of Chemistry/Ministry of Agriculture Food and Fisheries: London, U.K., 1991.

Miller, J. C.; Miller, J. N. In *Statistics for Analytical Chemistry;* Ellis Horwood: Chichester, U.K., 1993; pp 115−117.

Received for review November 3, 1998. Revised manuscript received June 9, 1999. Accepted June 11, 1999. We thank the Biotechnology and Biological Sciences Research Council and the Islamic Development Bank for supporting this work.

JF981196D