TITLE: Prediction of naturally-occurring, industrially-induced and total trans fatty acids in butter, dairy spreads and Cheddar cheese using vibrational spectroscopy and multivariate data analysis

AUTHORS: Ming Zhao, Renwick J. Beattie, Anna M. Fearon, Colm P. O'Donnell, Gerard Downey

NOTICE: This is the author's version of a work that was accepted for publication in *International Dairy Journal.* Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in *International Dairy Journal*, Volume 51, December 2015, Pages 41-5. DOI 10.1016/j.idairyj.2015.07.011

Prediction of naturally-occurring, industrially-induced and total *trans* fatty acids in butter, dairy spreads and Cheddar cheese using vibrational spectroscopy and multivariate data analysis

Ming Zhao [a b] , Renwick J. Beattie [c] , Anna M. Fearon [c], Colm P. O'Donnell [b] , Gerard Downey* [a b]

[a] Teagasc Food Research Centre Ashtown, Dublin 15, Ireland

[b] School of Biosystems Engineering, University College Dublin, Belfield, Dublin 4, Ireland

[c] Agri-Food and Biosciences Institute, Newforge Lane, Belfast, UK

* Corresponding author. Tel: +353 1 8059500; fax: +353 1 8059533. E-mail address: gerard.downey@ teagasc.ie

Abstract

This study investigated the use of vibrational spectroscopy (i.e. near infrared (NIR), Fourier-transform mid infrared (FT-MIR), Raman) and multivariate data analysis for (1) quantifying total *trans* fatty acids (TT) and (2) separately predicting naturally-occurring (NT; i.e. C16:1 t9; C18:1 *trans*-n, n=6…9, 10, 11; C18:2 *trans*) and industrially-induced TFAs (IT=TT-NT) in Irish dairy products i.e. butter (n=60), Cheddar cheese (n=44), dairy spreads (n=54). Results of gas-chromatography (GC) analysis were used as the chemical reference values. Butter, dairy spreads and Cheddar cheese separately contained TT around 4.42%, 0.70% and 3.04% of their total fatty acid methyl ester (%w/w FAME). Butter contained about 4.22% (NT) and 0.21% (IT) while those in Cheddar cheese averaged around 2.89% and 0.15%; corresponding values for dairy spreads were approximately 0.62% and 0.07%. Partial least squares regressions (PLSR) models for predicting NT, IT and TT in each type of dairy product were developed respectively using FT-MIR (900-1800 cm^{-1}), NIR (1700-1800 nm) and Raman (900-1798 cm^{-1}) spectral data. Models based on NIR, FT-MIR and Raman spectra can be used for the prediction of NT and TT content in butter; best predictive performance achieved the coefficient determination of validation (R^2V)~0.91-0.95, root mean square error

prediction (RMSEP)~0.07-0.30 for NT; R^2V~0.92-0.95, RMSEP~0.23-0.29 for TT. Prediction of NT, IT and TT in Cheddar cheese was with low accuracy due to interference from protein; in the case of dairy spreads and butter, prediction of IT was inaccurate due to the limitation of vibrational spectroscopy in predicting very low analyte contents (< 0.2% w/w of TFA).

Keywords: Vibrational spectroscopy; Total *trans* fatty acids (TT); Naturally-occurring *trans* fatty acids (NT); Industrially-induced *trans* fatty acids (IT); Multivariate data analysis, Dairy products

1. Introduction

In Western Europe, annual dairy product consumption contributes up to 30% of individual intake of animal fat (FAO, 2013). Fat in dairy products originates in raw bovine milk contains which contains about 3.5% milk fat; it typically consists of saturated fatty acids (SFA; 70% $^{w}/_{w}$), monounsaturated fatty acids (MUFA; 25% $^{w}/_{w}$) and polyunsaturated fatty acids (PUFA; 5% $^{w}/_{w}$) (Grummer, 1991). Unsaturated fatty acids consist of *cis / trans* double bond isomers and conjugated double bond isomers. *Trans* fatty acids (TFAs) have been defined as "all the geometrical isomers of monounsaturated and polyunsaturated fatty acids having non-conjugated, interrupted by at least one methylene group, carbon-carbon double bonds in the *trans* configuration" (Codex Alimentarius, 1985). *Trans* PUFA "have at least one *trans* double bond and may therefore also have double bonds in the *cis* configuration" (EFSA, 2010). *Trans* fatty acids with conjugated double bonds (not separated by a methylene group) are excluded by these definitions (FDA, 2003) and hence conjugated linoleic acid

(CLA) isomers which occur naturally in bovine milk are not included in the TFA category as defined by these agencies.

Consumption of TFAs has been linked with increased risk of coronary heart disease by raising levels of low density lipoprotein (LDL) cholesterol and lowering levels of high density lipoprotein (HDL) cholesterol (Mensink et al., 2003; Stender et al., 2006). Consequently, consumers and the dairy industry have awareness of and demand for information on the TFA content of dairy products and other processed foods. The FDA (2006) mandated the declaration of the total *trans* fat content on the Nutrition Facts label when a product contained ≥0.5 g of *trans* fatty acids per serving, corresponding to a content of approximately 2% w/w of total fat.

Trans fatty acids in foods arise either as naturally-occurring or industrially-induced compounds. Industrially-induced TFAs (IT) are mainly formed during the partial hydrogenation or deodorisation of oils or fats during refining or cooking (Kodali, 2005). Naturally-occurring TFAs in dairy products include *trans* isomers of C18:1, C18:2, C18:3 and *trans*-palmitoleic acid (C16:1 t9), a biomarker of dairy consumption. These TFAs arise in the rumen of ruminants during extensive hydrolysis and bio-hydrogenation of unsaturated fatty acids contained in feed by the rumen bacteria (IDF, 2013). Industrial hydrogenation of oils and fats basically adds hydrogen to unsaturated bonds of oils and fats to produce a range of hydrogenated and partially hydrogenated products with modified physical properties and oxidative stability. Industrially-induced TFAs present in dairy products have generally been added in the form of partially hydrogenated oils which have been included in order to manipulate the physical properties of the dairy product (Wright, Scanlon, Hartel & Marangoni, 2001; Fearon, 2011).

Some controlled dietary intervention studies indicate that TFAs from industrial and natural sources have different effects on cardiovascular risk factors in human; a high dietary intake of TFAs from natural sources was found to have a neutral effect on plasma lipid and other cardiovascular disease risk factors whereas intakes of industrially-induced TFAs seem to have a specific HDL cholesterol-lowering property (Chardigny et al., 2008; Bélanger et al., 2008). Based on these observations, quantification of total TFAs and separate identification of naturally-occurring TFAs have become important research tasks to assist in meeting both labelling and consumer information requirements.

The most accurate and common method used for TFA identification and quantification is gas-chromotography (GC) analysis. This method requires extraction of fat from the food matrix and preparation of the fatty acids as easily volatilised methyl esters, thus destroying the sample and requiring significant time and resources. Therefore, some vibrational spectroscopic techniques (NIR, MIR and Raman) have been explored to provide non-destructive and rapid measurements in this application. Vibrational spectroscopy is based on quantum energy shifts when incident photons excite specific chemical bonds in sample material under reflection/transmission (IR) and anti-stokes/stokes scatteration (Raman) modes. Since the 1940s, IR spectroscopy has been used to determine *trans* content in edible fats and oils (AOCS, 1946; Firestone and LaBouliere, 1965; Dutton, 1974; Sleeter and Matlock, 1989). The American Oil Chemists' Society (AOCS) published official IR methods in 1999, 2000 and 2009 (AOCS, 1999, 2000, 2009) but these methods are only suitable for edible oils and fats which are typically composed of a relatively small number of different fatty acids (approximately several dozen) compared to the over 400 in milk fat; they haven't achieved accurate prediction at low concentrations (<5% w/w of total fat) in complex

matrices such as milk fat (Mossoba et al., 2011; Stefanov et al., 2011). Dispersive MIR spectroscopy has previously been investigated for the prediction of the fatty acid composition of bovine milk (Soyeurt et al., 2006, 2011); NIR was demonstrated for the same purpose in bovine milk (Coppa et al., 2010) and in some specific cheese varieties (Lucas et al., 2008). Recently the use of MIR and NIR instruments for the prediction of fatty acids in fresh and thawed bovine milk has been reported (Coppa et al., 2014). FT-MIR has also been used to determine fatty acid composition of butter and margarines (Safar et al., 1994) while FT-Raman has been reported to detect TFAs in a study limited to raw milk samples only (Stefanov et al., 2011). No published reports have described the investigation of spectroscopic methods to quantify naturally-occurring and industrially-induced TFAs in dairy products.

The objectives of this study were to investigate the use of vibrational spectroscopy (i.e. NIR, FT-MIR and Raman) and multivariate data analysis for (1) quantifying TT and (2) separately predicting NT and IT in the selected Irish dairy products (i.e. butter, dairy spreads and Cheddar cheese).

2. Materials and methods

2.1 Samples

In total, 60 butter, 54 dairy spreads and 41 Cheddar cheese samples were collected at regular intervals over the course of 1 year (04/07/2012- 05/07/2013). Samples of butter and Cheddar cheese were supplied by a well-known Irish manufacturer; dairy spread samples from 4 major Irish brand names (Dairygold (Kerry Foods Ltd.), Kerrymaid (Kerry Foods Ltd.), Flora

(Unilever, Ireland) and Lowlow (Kerry Foods Ltd.)) and were purchased at approximately 2 week intervals in different retail outlets in the Dublin city area. Fresh samples were stored at -4 ºC for less than 2 days before subsampling into screw cap containers (butters and dairy spreads) or aluminium foil (cheese) prior to storage at -20 °C.

2.2 Determination of *trans* fatty acids by GC

2.2.1 Fat extraction

Butter and dairy spreads are mainly fat-based while cheese has a much higher protein content, therefore different methods of fatty acid extraction were performed on these two sample groups. The fat and non-fat portions of the butters and dairy spreads were separated by first melting in an oven at 40 ºC and then centrifuging at 3200 rpm (1591 g) at 15 ºC for 10 mins (Versatile centrifuge, Sigma 2-16 PK; Sigma, Osterode am Harz, Germany). The top fat layer was then removed and methylated as below. Cheese samples were grated, mixed with ammonium thiocyanate (30% w/v) and then shaken in a thermal shaker at 60 ºC for 1 hour at 135 rpm before centrifugation at 1900 rpm (561 g) for 20 mins at 30 ºC. The upper fat layer was removed and methylated as below.

2.2.2 Methylation

Extracted fat samples were melted at 60 °C and an aliquot (40 µl) of each sample transferred into a glass tube; 5 ml n-hexane and 2 M KOH solution (200 µl) were then added and mixed using a vortex mixer for 40 seconds. Tubes were left on the bench at ambient temperature for 5 mins after which sodium hydrogen sulphate monohydrate (0.5 g) was added to each tube before centrifugation at 1200 rpm (223 gravity) for 5 mins at 15 ºC. Afterwards, 120 µl of the

methylated sample and 780 µl of n-hexane were transferred into a 2 ml vial prior to GC analysis. All chemicals used were from Sigma, Osterode am Harz, Germany.

2.2.3 GC analysis

Methylated samples were analysed following injection onto a gas-liquid chromatograph (Varian 3400 Capillary GC; Varian, Walnut Creek, CA., USA) fitted with a flame-ionisation detector and a CP Sil 88 capillary column (100 m * 0.25 mm * 0.20 µm; Chrompack, Middleburgh, The Netherlands). Separation of fatty acid methyl ester (FAME) was performed as previously described (Coakley et al., 2007). Each chromatograph comprised peaks in a plot of retention time (mins) versus voltage (volts) from which the area under each peak was expressed as a percentage of total fatty acid methyl ester (FAME) detected. Data collection and integration were performed using Varian Star Chromatography Workstation (version 6.0, Varian Inc., CA., USA) and Compass CDS (version 2.0, Bruker, Amundsenweg, The Netherlands).

To ensure comparability of columns and accuracy of TFA results, reference materials were developed at Teagasc (Moorepark, Ireland) and regularly assessed during the analysis. TFAs were detected in the dairy products tested as follows: C14:1 t9, C15:1 t10, C16:1 t9, C17:1 t10, C18:1 t6-t9, C19:1 t7, C19:1 t10, C18:2 t9 t12, C18:2 t,c; c,t, C20:1 t11 and C22:1 t13. TFA standards were purchased in methylated form (Nu-chekprep.Inc, Elysian, MN, USA). Analysis of each TFA was performed three times on separate occasions to determine the average retention time of each TFA. Polyunsaturated TFAs (C18:2 t,c ; c,t) were detected between the retention times of C18:2 t, t and C18:2 c,c (linoleic acid) on the chromotograph. All the peaks in this retention time range were taken to be a C18:2 t,c ; c,t group. The entire peak area of C18:1 t6 - t11 was categorised as the C18:1 *trans* group because further

temperature programming was able to show overlapping between *trans*-n; *trans*-n and *cis*-n of C18:1 isomers. Samples were analysed in random order; one out of every 3 or 4 samples was analysed in duplicate in order to determine the consistency of the results.

Naturally-occurring TFAs (NT) include the C18:1 *trans* group (C18:1 *trans*-n n = 6…9, 10, 11), C18:2 *trans* group and C16:1 t9. Trace amounts of naturally-occurring TFAs in the C18:1 *trans* group (n= 12, 13…16) and C18:3 *trans* group have not been included in this study due to the separation limits and detection thresholds of the GC analysis used. All fatty acids detected in this work which are not conventionally accepted as naturally-occurring TFAs are referred to as industrially-induced TFAs. All GC chemical values of C18:1 *trans* group (6-9, 10, 11), C18:2 *trans* group and C16:1 t9 were summed as NT; other TFA chemical values together were counted as IT meaning that TT measured here could be summarised as TT=NT+IT.

2.3 Vibrational spectroscopic measurements

Frozen cheddar cheese samples were allowed to thaw at ambient temperature for 17-18 hours and butter and dairy spread samples were left in a forced air oven (Memmert, Schwabach, Germany) at 30 ºC for 30 minutes to thaw before any spectroscopic analysis was carried out.

2.3.1 NIR spectroscopy

Near infrared analyses were performed using a NIRSystems 6500 visible/NIR scanning spectrometer (NIRSystems, Inc., Laurel, MD, USA) in reflectance mode. Samples were scanned over the 400-2498 nm wavelenth range at 2 nm with a 16-32 reference sample scan sequence. Approximately 8 g finely grated cheddar cheese at ambient tempertature (~20 ºC)

was placed into a circular ring cup (3.8 cm inner diameter) equipped with a disposable backing disc. Butter and dairy spread samples (~1.5g) at 30 ºC were evenly spread on the glass window of a camlock cell and compressed with a gold-plated metal reflector (0.2 mm sample thickness). Between samples, sample holders were carefully cleaned using tepid water with detergent and dried. Each sample of each dairy product type (butter, dairy spreads and Cheddar cheese) was scaned twice by rotating the same sample cup or cell through a 120-degree angle between sequential scans. Instrument control, spectral collection and conversion to JCAMP.DX format (Rutledge and McIntyre, 1992) was performed using WINISI software (version 1.04; Intrasoft International, State College, PA, USA). The mean of duplicate scan results was used in subsequent chemometric operations.

2.3.2 FT-MIR spectroscopy

For FT-MIR spectroscopic analysis of the dairy products, a Fourier transform mid-infrared spectrometer (Bio-Rad Excalibur FTS3100; Bio-Rad, Philadelphia, USA) with an attenuated total reflectance (ATR) accessory (single bounce crystal; PIKE MIRacle, PIKE Technologies, Madison, Wisconsin, U.S.A.) was used. The ATR accessory was equipped with a zinc selenide crystal and a temperature controller (PIKE TempPRO, PIKE Technologies, Madison, Wisconsin, U.S.A.). ATR spectra were recorded over the frequency range 698-4000 cm^{-1} at a resolution of 4 cm^{-1}; 64 sample scans were averaged and corrected using an air blank reference scan recorded before each sample. Spectral data were recorded and converted into Grams (.spc) format using manufacturer's software (Resolution Pro v 4.0).

Surface portions of each cheese sample were removed and discarded and the central portion cut into small cubes (0.5 cm*1.5 cm*2 cm). During spectroscopic measurements, a cheese sample cube was presented as one face (1.5 cm*2 cm) which was pressed against the entire suface of the ATR crystal with a 100 g weight. Butter or dairy spread samples (~ 0.2 g) were

applied uniformly over the preheated crystal; and scanned in duplicate (with sub-sampling) at 30 ºC. Before each spectral acquisition, the ATR crystal was cleaned with acetone and allowed to air dry.

2.3.3 Raman spectroscopy

Raman spectra were collected over the 200-2000 cm^{-1} frequency range using a DeltaNu Rapid –ID dispersive Raman spectrometer (DeltaNu, Laramie, Wy, USA) equipped with a 785 nm near infrared laser. DeltaNu NuSpec software (v. 6.0; Perkin Elmer Inc., USA) was used for instrument management, spectral acquisition and file conversion. The DeltaNu system automatically applies a wavelet transformation and spline interpolation on all data so no further smoothing was performed.

Cheese samples were taken using a corer (1.25 cm diameter) and placed into the empty vial sampling attachment. Butter and dairy spread samples were placed in plastic tubes with 2 mm diameter holes drilled for laser access. This ensured that no material from the sample holder was likely to interfere with the Raman signal. Five spectra were measured for each sample at physically separate locations, with spectra recorded for 5 s x 12 accumulations and saved as a *.spc multifile.

2.4 Chemometrics and statistics

The spectra of FT-MIR and NIR were separately exported from Resolutions Pro software in GRAMS format and WINISI software as JCAMP.DX files, then directly imported into The Unscrambler software (v.9.7; Camo, Trondheim, Norway). Raw spectra of each type of dairy product (butter, cheese, dairy spreads) over the full scan range were plotted and visually

examined to detect unusual samples. Principal component analysis (PCA) was performed using untreated data of the full spectral range to identify patterns and potential outliers in each sample type.

Delta Nu Raman spectra were saved as *.spc multifiles and imported into Matlab. A baseline correction and potential outliers detection were effected using the Singular Value Decomposition (SVD)-based background correction programme. GC data were imported into Matlab and placed into a data matrix alongside the processed Raman data. Normalisation factors were calculated and separately applied using the mean intensity of the carbonyl stretching mode (ca. 1740 cm^{-1}), the CH_x scissoring mode (ca. 1440 cm^{-1}) and linear combinations of the scores from the first 3 PCs (Beattie, et al., 2009). The PC-based normalisation results are presented as they gave the best performance. All data were then imported into The Unscrambler (v 9.7, Camo, Oslo, Norway).

To develop PLSR models, spectral data were used both in un-treated (raw) and pre-treated forms; pre-treatments studied for the NIR and FT-MIR data included normalisation on unit vector (nor. u. v.), multiplicative scatter correction (MSC), standard normal variate (SNV) transformation and Savitzky-Golay (1^{st} derivative with 9 points; 2^{nd} derivative with 17 points). Raman data were assessed untreated and pre-treated using a combination of SVD-based baseline correction (Beattie, 2011), normalisation to a peak intensity suitable for use as an internal standard (mean intensity of CH_2 scissor band at 1440 cm^{-1} (Beattie, 2004) and normalisation to the cumulative sums of scores over the first 3 PCs (Beattie, et al., 2009). Techniques such as MSC and SNV are not relevant for Raman data and so were not employed (Afseth, Segtnan and Wold, 2006; Beattie and McGarvey, 2013).

Partial least squares regression quantitative models for NT, IT and TT were developed separately for each sample type and for all samples using a number of spectral ranges decided by intensities of regression coefficients i.e. FT-MIR: 698-4000 cm^{-1}, 698-2500 cm^{-1}, 698-1800 cm^{-1}, 698-1500 cm^{-1}, 698-1200 cm^{-1}, 698-1000 cm^{-1}, 800-1200 cm^{-1}, 800-1500 cm^{-1}, 900-1800 cm^{-1};, NIR: 1100-2500 nm, 2200-2500 nm, 1300-1500 nm, 1700-1800 nm, 1150-1210 nm, 800-1500 nm, 800-1800 nm, 890-916 nm and Raman: 500-1898 cm^{-1}, 689-1500 cm^{-1}, 900-1898 cm^{-1} to detect the range with the best prediction performance.

Partial least squares regression models were developed and evaluated on separate calibration and validation sample sets; the same samples were in each set for each spectroscopic technique. Calibration samples were selected in a quasi-random manner or the basis of sampling dates and batches. The Martens uncertainty test was applied to improve PLSR models through the use of a reduced number of spectral variables selected as significant for each prediction (Martens and Martens, 2000). Full cross-validation (i.e. leave-one-out) was used to develop models using the calibration sample set; optimum model complexity was determined using the first local minimum in residual validation variance plots. Statistical parameters used to evaluate these models included the correlation coefficient of determination in cross-validation (R^2CV) and root mean square error of cross-validation (RMSECV) (Williams and Norris, 2001). Models were validated using a test set which comprised those samples not selected for calibration; relevant statistics include the coefficient of determination in external validation (R^2V), root mean square error of prediction (RMSEP) and bias.

In order to compare the estimated noise levels within the PLSR models calculated from the spectra generated by different techniques, the ‘jaggedness’ of the regression coefficients for

models prior to application of the Martens' uncertainty test was used. This was based on the relative bandwidth of noise features (each pixel is assumed to be independent in terms of shot noise, so the difference between adjacent pixels is random) and spectral features (the spectral intensity of a pixel depends on its position on the x-axis and is related to the pixels surrounding it so that spectral features change gradually over several pixels). To generate this estimate (which is only suitable for direct comparative purposes, not absolute calculations) the noise is estimated by subtracting the mean value of each pixel's adjacent neighbours from that of the current pixel, then averaging the magnitude of this difference. The signal-to-noise (or more correctly, jaggedness) ratio is then estimated by dividing the mean magnitude of the regression coefficients by the jaggedness value. The uncertainty test eliminates unselected variables and so creates a disjoint vector unsuitable for calculating jaggedness, hence the requirement to estimate noise in models prior to application of this test.

3. Results

3.1 TFA determination by GC

The range, mean, median and standard deviation value (% w/w of total FAME) of single TFAs and TFA groups (C18:1 *trans*, C18:2 *trans*, total TFAs, industrially-induced TFAs, naturally-occurring TFAs) in the three dairy product categories are shown in Table 1. GC peaks for C18:1 *trans*-6, *trans*-9, *trans*-10 and *trans*-11 merged into one peak and this peak was counted as C18:1 *trans*- group; other C18:1 *trans* isomers (i.e. *trans*-12-16) co-eluted with C18:1 *cis* isomers and were not recorded separately in this study. With regard to C18:2 *trans* fats, multiple peaks of these polyunsaturated *trans* fatty acids overlapped in our system and this peak area was counted as a C18:2 *trans* group.

Butter, dairy spreads and Cheddar cheese contained total TFAs around 4.42%, 0.70% and 3.04% respectively of their total fatty acid contents expressed as a % w/w of total FAME (Table 1). Butter contained about 4.22% of naturally-occurring and 0.21% industrially-induced TFAs while corresponding values in Cheddar cheese were about 2.89% and 0.15%; values for in dairy spreads were about 0.62% and 0.07% respectively. It is clear that TFAs in dairy products were mainly those which naturally occur in milk while industrially-induced TFAs accounted for less than 0.6% w/w of total fat.

3.2 Spectral results

Averaged FT-MIR, NIR and Raman spectra of each dairy product were plotted over a number of spectral ranges (Figure 1a – 1f). PCA of NIR, FT-MIR and SVD of Raman spectra performed respectively over their full spectral range did not reveal any unusual samples and the complete dataset was retained in the analysis.

The FT-MIR spectra over the wavelength range of 4000-698 cm^{-1} (Fig. 1 a) closely resemble spectra of milk and dairy products reported in earlier studies (Safar, 1994; Koca et al., 2010). Absorbance peaks in this wavenumber range have been ascribed to $–CH_2$ of aliphatic chains (~719 cm^{-1}), C=C stretching bonds and the C-H out-of-plane deformation bond of *trans* isomers (966 cm^{-1}), C-O stretching bonds of triacylglycerol ester (1161cm^{-1}) and other C-H bonds (~ 1230 cm^{-1}, 1370 cm^{-1}, 1417 cm^{-1} and 1463cm^{-1} (Safar, 1994). Peaks between 1500 and 1700 cm^{-1} originate in peptide bonds or O-H bond absorptions (Safar, 1994; Goormaghtigh, Cabiaux, & Ruysschaert, 1994). FT-MIR averaged spectra of the three dairy product types show very similar patterns over the selected wavelength ranges; one peak of the cheese spectra exists at 1547 cm^{-1} which may be related to peptide bonds (Fig.1a, b). A large

peak around 1744 cm^{-1} has been attributed to C=O absorption in lipids (Al-Jowder et al., 1994) while strong C-H absorption is also evident between 2800 and 3000 cm^{-1}, adjacent to regions containing absorptions by aliphatic compounds (Dufour, 2009). A small kink observed around 3000-3100 cm^{-1} (Fig. 1a) has been identified as C=C stretching of *cis* isomers in an aliphatic chain or aromatic ring (Safar, 1994) while the broad feature around 3200-3700 cm^{-1} arises from overtones of O-H and N-H stretching in mixed compounds with water and amino acids (Safar, 1994; Stuart, 2004).

NIR spectra shown in Figure 1 represent the full measured wavelength range (400-2500 nm) and a narrower range corresponding to the first overtone range of C-H absorbances in aliphatic hydrocarbons (1700-1800 nm). In general, the spectra of butter and dairy spreads showed absorbance peaks at almost the same wavelengths, while the NIR spectra of cheese showed one extra peak at 970 nm (Fig. 1c); this may be attributed to the second overtone of O-H absorbance and corresponds to the greater water content of this dairy product (approximately 37 % w/w). Important peaks at 1730 nm and 1764 nm were visible (Fig. 1d) while maxima at 1940 nm (a combination of O-H stretch and deformation) and 1450 nm (the first overtone of O-H stretch vibration) have been ascribed to water. These spectral features are in agreement with previously published reports (Blazquez et al., 2004; Coppa et al., 2010). Other peaks at 1208 nm (the second overtone of –CH stretch), 1728 and 1765 nm (the first overtone of –CH stretch), 2310-2345 nm (combined –CH stretch and deformation band) have been related to fat moieties (Osborne and Fearn, 1988). Protein absorbances often obscure NIR spectral bands associated with fatty acids (Afseth et al., 2005; Afseth et al., 2006), with the result that fatty acid prediction becomes problematic in the presence of protein.

Raman spectra are dominated by features attributable to fatty acids. Peaks at 1748, 1659, 1445, 1306, 1271, 1129 and 920 cm^{-1} (Fig. 1f) have been assigned to the C=O carbonyl stretch, C=C olefinic stretch, CH_2 scissoring or twisting bonds, C-H in-plane deformation bonds and C-C aliphatic stretching bonds of aliphatic chains in lipids (Socrates, 2000; Beattie et al., 2004). Only the cheese samples contain spectral features that can reasonably be attributed to molecules other than water and lipids, exhibiting a number of peaks that are characteristic of proteins. One peak at 1003 cm^{-1} is found only in the spectra of cheese and not those of butter and dairy spreads (Fig. 1e); this peak is strongly related to phenylalanine (Beattie et al., 2004). Bands around 1645-1685 cm^{-1} and 1200-1350 cm^{-1} have been related to the secondary and tertiary structure of proteins i.e. amide I and amide III bands respectively (Fig. 1f). The peak at 1659 nm is probably due to α-helical structures, those at 1271 and 1306 cm^{-1} have been attributed to $\alpha_{globular}$ and $\alpha_{fibrous}$-helix formations (Beattie et al., 2004) and the peak at 1445 cm^{-1} has been ascribed to a CH_2 scissoring vibration in proteins. Bands in the 1062-1089 cm^{-1} region have been reported as originating in the C-O stretching of aspartic and glutamic acids (Li-Chan, 1996).

3.3 PLSR prediction models

Three Y-outliers were detected among 60 butter samples while none were found in 41 Cheddar cheese or 54 dairy spread samples. As mentioned above, no outliers were detected using PCA on spectral data but the outliers detected by PLS analysis called as Y-outliers, defined as those observations that have a relationship between spectral (x) and chemical reference (y) data which differ from the other samples being measured (Næs, Isaksson, Fearn and Davies, 2002). The three butter sample outliers (i.e. butters #3, 47 and 49) were detected using the residual Y-variance; large Y-residual values for these samples (more than twice the

residual standard deviation for butter samples) indicated either errors in reference measurements or samples with a different relationship between x and y data (Martens and Næs, 1989). Therefore, these three butter samples were eliminated when developing PLSR models.

Performance of initial PLSR models for predictions of NT, IT and TT in Cheddar cheese and dairy spreads was not accurate for practical use ($R^2V < 0.65$). In the case of Cheddar cheese, protein is the dominant component and overlaps absorbances by fatty acids as stated above (Figure1). Consequently this overlapping may disrupt the correlation between the spectral data (X- values) and fatty acid reference data (Y- values) input to PLSR model. In the case of dairy spreads, the poor prediction results for *trans* fatty acid content (~0.7 % w/w FAME) most likely results from the low concentration range of these moieties in dairy spreads. Therefore, these poorly performing models will not be further discussed.

Summary results of the performance of PLSR models for butter analysis by FT-MIR (900-1800 cm^{-1}) , NIR (1700-1800 nm) and Raman shifts (900-1798 cm^{-1}) are shown in Tables 2, 3 and 4 respectively. The models cannot predict the IT content (~ 0.2 % w/w FAME) in butter because of its low concentration range, a general limitation of vibrational spectroscopy. In Table 2, the models developed using FT-MIR spectral data for NT and TT predictions reveal good performances in general with R^2CV equal to 0.91-0.93 for NT, 0.92-0.93 for TT; R^2V of 0.77-0.85 for NT, 0.82-0.84 for TT. Models developed using MSC followed by a first derivative (9 – point gap Savitzky-Golay filter) pre-treated data and 2 to 4 PLS loadings performed well for the prediction of NT and TT. In this case, the R^2 values (R^2CV and R^2V) of models to predict NT and TT were ≤ 0.85 with RMSE (RMSECV and RMSEP) values <0.4. Table 3 shows that models developed using NIR spectral data had better performance than models based on FT-IR spectral data with R^2CV of 0.86-0.91for NT, 0.88-0.91 for TT

and R^2V equal to 0.91-0.95 for NT, 0.92-0.95 for TT. The best performing NIR models were developed using data pre-treated by SNV and first derivative (9 – point gap Savitzky-Golay filter) using 9 PLS loadings resulting in R^2 values of NT and TT prediction models of ≤0.95 and corresponding RMSE values ≤0.31. The performance of models using Raman data was poorer than that of NIR and FT-IR models with. R^2CV equal to 0.71-0.79 for NT, 0.71-0.79 for TT and R^2V of 0.56-0.78 for NT, 0.54-0.77 for TT. The best performing Raman model was produced using baseline correction and normalisation on PC 1, 2 & 3 but still resulted in a poorer prediction performance for NT and TT (R^2≤0.78, RMSE≤0.52). Validation results of the best performing models in Tables 2, 3 and 4 are shown graphically in Figure 2.

4. Discussion

4.1 Regression coefficient weights of NIR, FT-MIR and Raman

Regresson coefficients of PLSR models were used to identify the most influential wavelength ranges for prediction. Figure 2 c shows the regression coefficients over the attenuated simplified fingerprint range (after the Martens uncertainty test, fewer wavelength variables will be selected as the simplified range) of NT and TT predictions in butter samples using FT-MIR spectral data. Intensity peaks for both NT and TT predictions occurred similarly at 946 cm^{-1}, 962-976 cm^{-1}, 1446-1502 cm^{-1}, 1584-1634 cm^{-1} and 1768-1778 cm^{-1} wavelength range, which are related to the absorption of C=C and C-H bonds out-of-plane *trans* isomers and other C-H bonds of aliphatic chain (Safar,1994). A peak was observed at 1194 cm^{-1} for NT although not for TT and this may be related to C-O bonds of triacylglycerol ester. A high intensity peak presents at 962 - 976 cm^{-1}; it is known that the 966 cm^{-1} C-H out-of-plane deformation band with *trans* isomers is uniquely characteristic of isolated (non-conjugated)

double bonds (AOAC, 2000; Mossoba, et al., 2011). The other obvious intensity change around 1114-1148 cm^{-1} is related to the C-O bond of esters and bending vibrations of methylene group (Koca, et al.,2010).

The strong intensity features within the Raman regression coefficients plots for NT and TT predictions (Fig. 2i) are located at 907, 1129, 1439-1468, 1646-1665 and 1743 cm^{-1} shift areas. Intensities of all these areas are typical of aliphatic structural bonds. The feature around 1129 cm^{-1} indicates C-C stretch bond in TFAs (Snyder et al., 1982), a feature around 1665 cm^{-1} is characteristic of isolated *trans* olefinic bonds, while a nearby feature selected by the uncertainty test was located at 1646 cm^{-1} and is a region typical of conjugated olefinic bonds in CLA (Susi et al., 1979). One intensity feature of NT prediction at 1297 cm^{-1} does not appear in the same position of the regression plot for TT prediction (Fig. 2i). This feature is related to CH_2 twisting deformations (Butler et al., 1979), which may represent the structural difference of NT from the general TT.

For the NIR models which performed best in this study, only 50 spectral variables in the 1700-1800 nm wavelength range were used for model development; therefore, it was not necessary to use Martens uncertainty test. In regression coefficient plots of butter samples for NT and TT predictions (Fig. 2f), peaks occurred around 1700 nm, 1720 nm, 1740 nm and 1760 nm, similar to the original absorbance spectra. Peaks around these wavelengths have been reported as absorption by C-H bonds of aliphatic chains (Burns and Ciurczak, 2001); however, proof for t*rans*-isomer absorption related to these specific wavelengths has not been presented.

4.2 Comparison of PLSR models developed for butter and diary spreads separately and in combination

Spectra of butter and dairy spreads (Figure 1) were similar and in an attempt to simplify the detection of TFAs in dairy products, it was investigated whether the butter and dairy spreads could be combined into a single model for prediction of TT content. For example, a discussion using FT-MIR spectral data is illustrated in Figure 3. The distribution of the total proportion of TFAs within the butter and dairy spread samples indicated that combining the two sample groups would provide a greater calibration range and the generation of a single combined model resulted in a much improved R^2V value of 0.855 (Fig.3d) compared to individual R^2V of 0.8 and 0.15 for the butters and dairy spreads respectively (Fig.3a and b). However, comparing the scatter plots for the dual model approach (Fig.3c) with the single model approach, the scatter within the spreads looks much lower. If the validation results from the single model for these two types are assessed separately, the R^2V is in fact much worse within each sample type at 0.04 for dairy spreads and 0.47 for butter while the RMSEP is much higher 0.70 and 0.62 for dairy spreads and butter respectively compared with 0.20 and 0.35 as shown in the table in Fig.3. The use of a combined model superfically appeared to give better results when looking at the R^2V value alone, but when the prediction errors and assessment of the two sample types in isolation were examined they revealed that a single model was a much poorer option. Figure 4 shows the total TFA distributions of both butter and dairy spreads.The total *trans* value of each dairy product type here is in fact made up of a very different set of fatty acids (Figure 4) and so it is not surprising that molecular spectroscopic techniques are sensitive to subtle changes both in molecular structure and in sample matrix.

4.3 Comparisons of signal-to-noise ratio of NIR, FT-MIR and Raman

A final discussion of the comparative performance of these three vibrational spectroscopic methods may shed some light on the levels of predictive ability achieved by each technique. When examining the regression coefficients calculated prior to the application of Martens uncertainty test, it is possible to estimate a relative contribution of shot noise to the model as this noise varies from pixel to pixel, while spectral features in all of the technologies employed spread over several (typically >10) pixels. By measuring the mean difference between adjacent pixels and ratioing to the mean magnitude of the coefficients, there's a measure that can crudely compare the relative signal-to-noise ratio for each method. The estimated signal-to-noise ratio in the regression coefficients for each method was 63.4, 13.2 and 1.6 for the NIR, FT-MIR and Raman models respectively (Figure 5). The magnitudes of the TFA contents are relatively low and it may be that the interaction between the noise levels achievable by each technique is influencing the quality of the models that can be created.

5. Conclusions

This study investigated an approach to address labelling issues related to naturally-occuring, industrially-induced and total *trans* fatty acids in selected Irish dairy products namely butter, Cheddar cheese and dairy spreads. The use of vibrational spectroscopy (NIR, FT-MIR and Raman) in conjunction with multivariate data analysis was explored for (1) the non-destructive quantification of total TFAs (TT) and (2) the separate prediction of naturally-

occurring (NT) and industrially-induced TFAs (IT=TT-NT). Chemical reference data values for NT, IT and TT were derived from *trans* fatty acid groups of the gas-chromatography results as classified by the IDF (IDF, 2013). PLSR models to predict NT and TT in butter using NIR and FT-MIR spectral data performed better those based on Raman spectra due to their lower signal-to-noise ratio. Given the extremely low values detected, all spectroscopic methods failed to predict IT content of butter and TFA contents of dairy spreads; no model could predict TFAs in Cheddar cheese as spectral infomation in this product was dominated by protein. Models developed for the individual dairy product types performed better for the prediction of *trans* fats than those developed using combinations of different dairy product types.

Acknowledgements

This project was funded by the Irish Department of Agriculture, Food and the Marine as part of CheeseBoard 2015. Thanks to Dr. Catherine Stanton, Dr. Paul Simpson, Dr. Mariead Coakley from Teagasc Moorepark, Ireland and Mr. Alan Beattie from AFBI, Belfast, UK for providing gas chromotography equipments and technical support. Ming Zhao is a Teagasc Walsh Fellow.

References:

AOCS. (1946). Tentative method Cd 14-61. Official and Tentative Methods of the American Oil Chemist's Society (4th ed.), AOCS, Chicago, USA.

AOCS. (1999). Official Method Cd 14d-99. Rapid determination of isolated *trans* geometric isomers in fats and oils by attenuated total reflection infrared spectroscopy (Revised 1999) AOCS, Champaign, IL, USA.

AOAC. (2000). Official Method of Analysis, method no. 2000.10. Determination of total isolated *trans* unsaturated fatty acids in fats and oils by ATR-FTIR (17th ed.). Gaithersburg, MD, USA: AOAC International.

AOCS. (2009). Official Method Cd 14e-09. Negative second derivative infrared spectroscopic method for rapid (5 min) determination of total isolated *trans* fat. AOCS, Champaign, IL, USA.

Afseth, N. K., Segtnan, V. H., Wold, J. P. (2006) Raman spectra of biological samples: A study of preprocessing methods, *Applied Spectroscopy*, 60, 12, 1358.

Blazquez, C., Downey, G., O'Donnell, C., O'Callaghan, D., Howard, V. (2004). Prediction of moisture, fat and inorganic salts in processed cheese by near infrared reflectance spectroscopy and multivariate data analysis. *Journal of Near Infrared Spectroscopy*, 12, 149-157.

Beattie, J. R., Bell, S. E. J., Borgaard, C., Fearon, A. M., and Moss, B. W. (2004). Multivariate prediction of clarified butter composition using Raman spectroscopy, *Lipids*, 39, 897-906.

Beattie, J. R., Bell, S.J., Farmer, L. J., Moss, B. W., and Desmond, P. D. (2004). Preliminary investigation of the application of Raman spectroscopy to the prediction of the sensory quality of beef silverside, *Meat Science*, 66, 903-913.

Beattie, J.R., Glenn, J. V., Boulton, M. E., Stitt, A. W., and McGarvey, J. J. (2009). Effect of signal intensity normalisation on the multivariate analysis of spectral data in complex 'real-world' datasets, *Journal of Raman Spectroscopy*, 40, 429-435.

Beattie, J. R. (2011). Optimising reproducibility in low quality signals without smoothing; an alternative paradigm for signal processing, *Journal of Raman Spectroscopy*, 42, 1419-1427.

Beattie, J. R., McGarvey, J. J. (2013). Estimation of signal backgrounds on multivariate loadings improves model generation in face of complex variation in backgrounds and constituents, *Journal of Raman Spectroscopy*, 44, 329-338.

Butler, M., Salem, N., Hoss, W. & Spoonhower, J. (1979). Raman spectral analysis of the 1300 cm^{-1} region for lipid and membrane studies, *Chemistry and Physics of Lipids*, 29, 99-102.

Codex Alimentarius Commission (1985). Guidelines on nutritional labelling (*CAC/GL2*).

Coakley, M., Barrett, E., Murphy, J. J., Ross, R. P., Devery, R., Stanton, C. (2007). Cheese manufacture with milk with elevated conjugated linoleic acid levels caused by dietary manipulation. *Journal of Dairy Science*, 90, 2919-2927.

Chardigny, J.-M., Destaillats, F., Malpuech-Brugère, C., Moulin, J., Bauman, D. E., Lock, A. L., Sébédio, J.-L. (2008). Do trans fatty acids from industrially produced sources and from natural sources have the same effect on cardiovascular disease risk factors in healthy subjects? Results of the *trans* Fatty Acids Collaboration (TRANSFACT) study. *The American Journal of Clinical Nutrition*, 87(3), 558-566.

Coppa, M., Ferlay, A., Leroux, C., Jestin, M., Chilliard, Y., Martin, B. & Andueza, D. (2010). Prediction of milk fatty acid composition by near infrared reflectance spectroscopy. *International Dairy Journal*, 20(3), 182-189. doi:http://dx.doi.org/10.1016/j.idairyj.2009.11.003.

Coppa, M., Revello-Chion, A., Giaccone, D., Ferlay, A., Tabacco, E., & Borreani, G. (2014). Comparison of near and medium infrared spectroscopy to predict fatty acid composition on fresh and thawed milk. *Food Chemistry*, 150(1), 49-57.

Donnellan, T., Hennessy, T., Keane, M., & Thorne., F. (2011). Study of the international competitiveness of the Irish dairy sector at Farm Level. Teagasc, Ireland. Retrieved insert the date from www.teagasc.ie/.../CompetitivenessofMilkProductionweb230611.pdf.

Dutton, H. J. (1974). Analysis and monitoring of *trans*-isomerization by IR attenuated total reflectance spectrophotometry. *Journal of the American Oil Chemists' Society, 51, 47-51.*

Dufour, E. (2009). Principles of Infrared Spectroscopy. In D. -W. Sun (Ed.), *Infrared Spectroscopy for Food Quality Control* (pp. 3–28). New York: Academic Press.

European Food Safety Authority (EFSA) (2010). Scientific opinion on dietary reference values for fats, including saturated fatty acids, polyunsaturated fatty acids, monounsaturated fatty acids, *trans* fatty acids, and cholesterol. Retrieved insert the data from www.efsa.europa.eu/en/scdocs/scdoc/1461.htm.

Fearon, A.M. (2011). Butter and Butter products pp199-224, *In* Dairy Ingredients for Food Processing, R.C. Chandan and A Kilra (Eds), Wiley - Blackwell Iowa, USA.

FAO (Food and Agriculture Organization of the United Nations) (2013). Statistical dataset: http://faostat.fao.org/site/610/DesktopDefault.aspx?PageID=610#ancor.

FDA (U.S. Food and Drug Administration) (2003). Fed. Regist. (July 11) 68, 41433.

Firestone, D., De La Luz Villadelmar, M. (1965). Determination of isolated *trans* unsaturation by infrared spectrophotometry. *Journal - Association of Official Analytical Chemists*, 48, 437-443.

Firestone, D. (1992). Fats and Oils. *Journal of AOAC International*, 75(1), 109-112.

Grummer, R. R. (1991). Effect of Feed on the Composition of Milk-Fat. *Journal of Dairy Science*, 74(9), 3244-3257.

IDF (International Dairy Federation) (2013). Questions on *Trans* and CLAs of Dairy products. Paris, France: International Dairy Federation. Retrieved insert the date from http://www.idfdairynutrition.org/Files/media/FactSheetsConsumers/Question-on-Trans-and-CLA-of-Dairy-Product.pdf.

Kodali, D.R. (2005). *Trans* fats – chemistry, occurrence, functional need in foods and potential solutions. In D.R. Kodali and G.R. List (Ed.), *Trans Fats Alternatives* (pp.1-25), AOCS Press, Champaign, IL, USA.

Koca, N., Kocaoglu-Vurma, N. A., Harper, W. J., & Rodriguez-Saona, L. E. (2010). Application of temperature-controlled attenuated total reflectance-mid-infrared (ATR-MIR) spectroscopy for rapid estimation of butter adulteration. *Food Chemistry*, 121(3), 778-782. doi: http://dx.doi.org/10.1016/j.foodchem.2009.12.083

Li-Chan, E. C. Y. (1996). The applications of Raman spectroscopy in food science. *Trends in Food Science & Technology*, 7, 361-370.

Lucas, A., Andueza, D., Ferlay, A., & Martin, B. (2008). Prediction of fatty acid composition of fresh and freeze-dried cheeses by visible–near-infrared reflectance spectroscopy. *International Dairy Journal*, 18(6), 595-604. doi: http://dx.doi.org/10.1016/j.idairyj.2007.12.001

Mensink, R. P., Zock, P. L., Kester, A. D., & Katan, M. B. (2003). Effects of dietary fatty acids and carbohydrates on the ratio of serum total to HDL cholesterol and on serum lipids and apolipoproteins: a meta-analysis of 60 controlled trials. *The American Journal of Clinical Nutrition*, 77(5), 1146-1155.

Mossoba, M. M., Seiler, A., Steinhart, H., Kramer, J. K. G., Rodrigues-Saona, L., Griffith, A. P., Pierceall, R., van de Voort, F. R., Sedman, J., Ismail, A. A., Barr, D., Da Costa Filho, P. A., Li, H., Zhang, Y., Bradley, M. (2011). Regulatory Infrared Spectroscopic Method for the Rapid Determination of Total Isolated Trans Fat: A Collaborative Study. *Journal of the American Oil Chemists Society*, 88(1), 39-46. doi: DOI 10.1007/s11746-010-1648-0

Motard-Bélanger, A., Charest, A., Grenier, G., Paquin, P., Chouinard, Y., Lemieux, S., Couture, P., Lamarche, B. (2008). Study of the effect of *trans* fatty acids from ruminants on blood lipids and other risk factors for cardiovascular disease. *The American journal of clinical nutrition*, 87(3), 593-599.

Martens, H., & Martens, M. (2000). Modified jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR). *Food Quality and Preference*, 11, 5–16.

Marten, H., Næs, T. (1989). *Multivariate calibration*. Chichester, West Sussex, UK: John Wiley & Sons.

Næs, T., Isaksson, T., Fearn, T., Davies, T. (2002). *A user-friendly guide to multivariate calibration and classification*. Pp.177-190. Chichester, West Sussex, UK: NIR publications.

Goormaghtigh, E., Cabiaux, V., Ruysschaert, J. M. (1994). Physicochemical methods in the study of biomembranes. In H. J. Hilderson, & G. B. Ralston (Eds.), *Subcellular Biochemistry*, Vol. 23. (pp. 329–362)New York: Plenum Press.

Osborne, B. G., Fearn, T. (1988). *Near infrared spectroscopy in food analysis*. Essex, UK: Longman Scientific & Technical.

Soyeurt, H., Dardenne, P., Dehareng, F., Lognay, G., Veselko, D., Marlier, M., Bertozzi, C., Mayeres, P., Gengler, N. (2006). Estimating fatty acid content in cow milk using mid-infrared spectrometry. *Journal of Dairy Science*, 89(9), 3690-3695.

Soyeurt, H., Dehareng, F., Gengler, N., McParland, S., Wall, E., Berry, D. P., Coffey, M., Dardenne, P. (2011). Mid-infrared prediction of bovine milk fatty acids across multiple breeds, production systems, and countries. *Journal of Dairy Science*, 94(4), 1657-1667. doi: 10.3168/jds.2010-3408

Stefanov, I., Baeten, V., Abbas, O., Colman, E., Vlaeminck, B., De Baets, B., & Fievez, V. (2011). Determining Milk Isolated and Conjugated *trans*-Unsaturated Fatty Acids Using Fourier Transform Raman Spectroscopy. *Journal of Agricultural and Food Chemistry*, 59(24), 12771-12783. doi: Doi 10.1021/Jf203463r

Stuart, B. H. (2004). *Infrared Spectroscopy: Fundamentals and Applications*. Chichester, West Sussex, UK: John Wiley & Sons, Ltd.

Sleeter, R. T., Matlock, M. G. (1989). Automated quantitative analysis of isolated (nonconjugated) *trans* isomers using Fourier transform infrared spectroscopy in corporating improvements in the procedure. *Journal - Association of Official Analytical Chemists*, 66, 121-127.

Safar, M., Bertrand, D., Robert, P., Devaux, M. F., Genot, C. (1994). Characterization of edible oils, butters and margarines by Fourier transform infrared spectroscopy with attenuated total reflectance. *Journal of the American Oil Chemists' Society*, 71, 371-377.

Snyder, R. G., Cameron, D. G., Casal, H. L., Compton, D. A. C., Mantsch, H. H. (1982). Studies on determining conformational order in *N*- Alkanes and phospholipids from the 1130 cm^{-1} Raman band. *Biochimica et Biophysica Acta*, 684, 111-116.

Susi, H., Sampugna, J., Hampson, J. W., Ard, J. S. (1979). Laser-Raman investigation of phospholipid – polypeptide interactions in model membranes. *Biochemistry*, 18, 297-301.

Wright, A.J., Scanlon, M.G., Hartel, R.W., Marangoni, A.G. (2001). Rheological properties of milk fat and butter. *Journal of Food Science*, 66, 1052-1071.

Table 1. *Trans* fatty acids expressed as % w/w of total fatty acids in dairy products determined using GC analysis.

Fatty acid	Butter (n=60)				Dairy spreads (n=54)				Cheddar cheese (n=41)			
	Range	Mean	Median	SD[a]	Range	Mean	Median	SD[a]	Range	Mean	Median	SD[a]
14:1 *trans*-9 (IT[b])	0.007-0.270	0.02	0.015	0.032	0.011-0.061	0.03	0.027	0.011	0.01-0.056	0.02	0.016	0.01
15:1 *trans*-10 (IT)	0.001-0.261	0.039	0.03	0.045	0.003-0.035	0.014	0.014	0.006	0.015-0.269	0.044	0.033	0.051
16:1 *trans*-9 (NT[c])	0.022-0.437	0.214	0.218	0.062	0.005-0.071	0.029	0.028	0.013	0.072-0.737	0.192	0.184	0.107
17:1 *trans*-10 (IT)	0.001-0.171	0.026	0.01	0.041	0.002-0.080	0.015	0.004	0.024	0.007-0.024	0.013	0.01	0.005
∑18:1 *trans*-6-11 (NT)	0.915-4.995	3.274	3.514	0.912	0.033-0.817	0.337	0.309	0.203	0.565-4.891	2.282	1.597	1.38
18:2 *trans*-9 trans-12 (NT)	0.004-0.461	0.068	0.04	0.08	0.004-0.042	0.023	0.024	0.009	0.015-0.16	0.059	0.04	0.073
∑18:2 t,c; c,t (NT)	0.22-1.255	0.642	0.653	0.194	0.048-0.485	0.273	0.266	0.09	0.056-1.089	0.473	0.364	0.314
19:1 *trans*-7 (IT)	0.023-0.253	0.074	0.061	0.045	0.004-0.044	0.013	0.009	0.01	0.021-0.209	0.088	0.083	0.039
19:1 *trans*-10 (IT)	0.02-0.202	0.048	0.041	0.031	0.001-0.146	0.021	0.018	0.024	0.009-0.197	0.064	0.043	0.076
20:1 *trans*-11 (IT)	0.001-0.158	0.029	0.013	0.041	0.015-0.084	0.034	0.031	0.014	0.003-0.073	0.03	0.027	0.017
22:1 *trans*-13 (IT)	0.007-0.026	0.014	0.014	0.004	0.001-0.008	0.005	0.005	0.005	–*	–	–	–
∑ NT[b]	**2.25-6.989**	**4.217**	**4.442**	**1.084**	**0.179-1.222**	**0.618**	**0.586**	**0.259**	**0.643-6.06**	**2.888**	**2.17**	**1.699**
∑ IT[c]	**0.103-0.471**	**0.205**	**0.176**	**0.079**	**0-0.23**	**0.069**	**0.059**	**0.037**	**0-0.595**	**0.147**	**0.145**	**0.103**
∑ TT[d]	**2.365-7.310**	**4.42**	**4.646**	**1.081**	**0.267-1.271**	**0.688**	**0.673**	**0.264**	**0.655-6.367**	**3.035**	**2.305**	**1.75**

[a] Standard deviation; [b] industrially-innduced TFAs; [c] naturally-occurring TFAs; [d] total TFAs; * below limit of detection; reference values used for model generation are in bold.

Table 2. Summary of PLSR model performances (FT-MIR – 900-1800 cm^{-1}) for NT, IT and TT in butter (most accurate models in bold).

Data type	Fatty acids	no. loadings	Calibration set n=40		Validation set n=17		
			R^2CV	RMSECV	R^2V	RMSEP	Bias
Raw	NT	5	0.91	0.30	0.81	0.45	0.03
	IT	1	0.07	-0.06	-0.05	0.07	0.02
	TT	5	0.92	0.29	0.82	0.43	0.04
MSC	NT	4	0.92	0.28	0.83	0.43	0.14
	IT	1	-0.07	0.07	-0.06	0.07	0.01
	TT	4	0.93	0.27	0.84	0.40	0.13
MSC+ 1st der.4+4	**NT**	**4**	**0.93**	**0.27**	**0.85**	**0.40**	**0.08**
	IT	**1**	**-0.08**	**0.07**	**-0.09**	**0.07**	**0.01**
	TT	**2**	**0.92**	**0.29**	**0.84**	**0.40**	**0.09**
SNV+1st der.4+4	NT	3	0.92	0.28	0.77	0.49	0.29
	IT	1	-0.08	0.07	-0.07	0.07	0.01
	TT	3	0.92	0.28	0.83	0.41	0.10

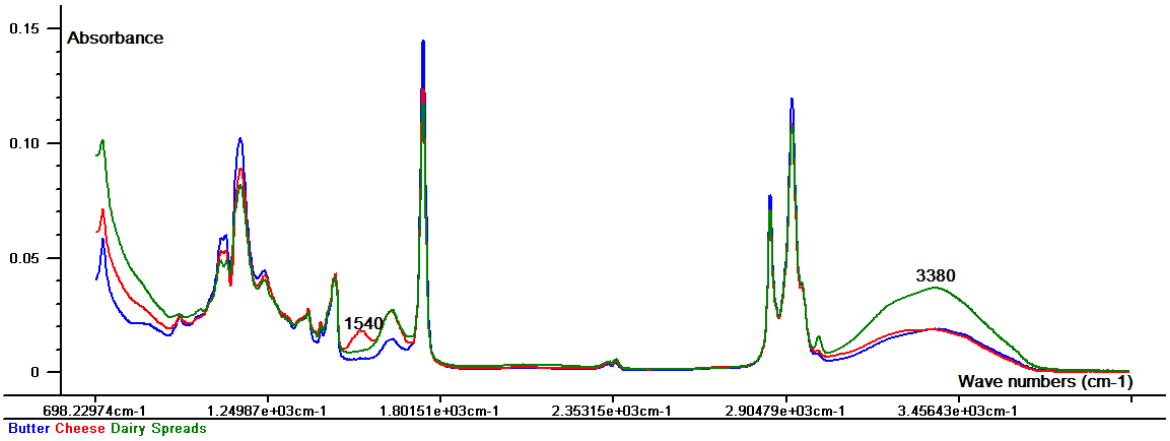
Table 3. Summary of PLSR model performances (NIR – 1700-1800 nm) for NT, IT and TT in butter (most accurate models in bold).

Data type	Fatty acids	no. loadings	Calibration set n=40		Validation set n=17		
			R^2CV	RMSECV	R^2V	RMSEP	Bias
raw	NT	9	0.86	0.37	0.91	0.30	-0.04
	IT	1	-0.04	0.07	-0.07	0.07	0.02
	TT	9	0.88	0.35	0.92	0.29	-0.02
MSC	NT	10	0.89	0.34	0.93	0.27	0
	IT	1	-0.04	0.07	-0.06	0.07	0.02
	TT	10	0.88	0.35	0.93	0.26	0.03
SNV	NT	9	0.90	0.32	0.93	0.26	0
	IT	1	-0.04	0.07	-0.06	0.07	0.02
	TT	9	0.89	0.34	0.94	0.25	0.02
SNV+1st der.4+4	**NT**	**9**	**0.91**	**0.30**	**0.95**	**0.24**	**0**
	IT	**1**	**-0.03**	**0.07**	**0**	**0.07**	**0.02**
	TT	**9**	**0.91**	**0.31**	**0.95**	**0.23**	**0.02**

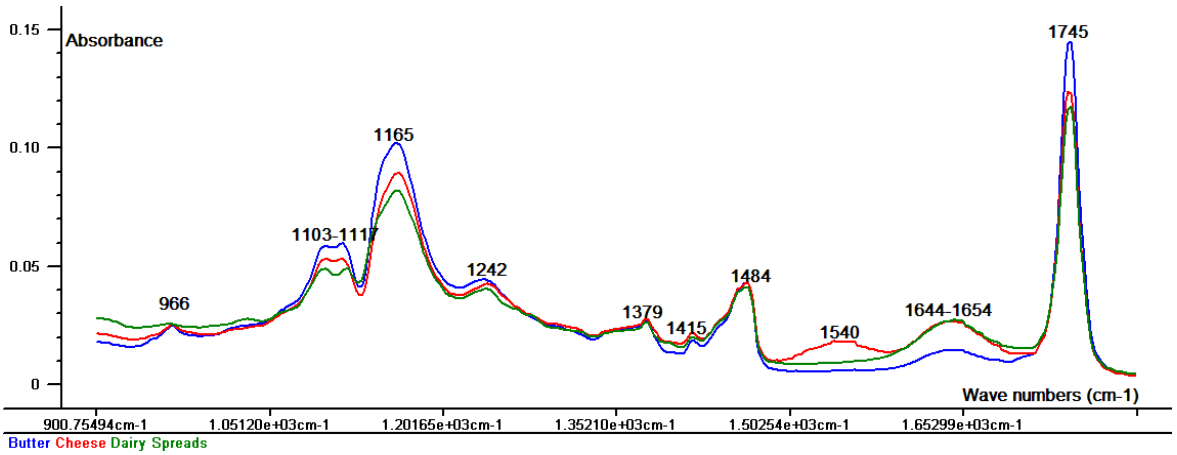
Table 4. Summary of PLSR model performances (Raman − 900-1798 cm^{-1}) for NT, IT and TT in butter (most accurate models in bold).

Data type	Fatty acids	no. loadings	Calibration set n=40		Validation set n=17		
			R^2	RMSECV	R^2V	RMSEP	Bias
raw	NT	6	0.79	0.46	0.56	0.68	0.12
	IT	1	-0.01	0.07	-0.28	0.08	0.01
	TT	6	0.77	0.48	0.54	0.68	0.16
Normalisation on unit vectors	NT	3	0.71	0.54	0.78	0.47	0.03
	IT	1	0.03	0.07	-0.42	0.08	0.03
	TT	6	0.79	0.46	0.75	0.43	0.07
Baseline Corrected + Normalisation on PC1, 2&3	**NT**	**2**	**0.73**	**0.52**	**0.78**	**0.48**	**0.03**
	IT	**1**	**0.16**	**0.07**	**-0.37**	**0.08**	**0.03**
	TT	**2**	**0.73**	**0.52**	**0.77**	**0.48**	**0.06**
Baseline Corrected + Normalisation on PC1	NT	2	0.71	0.54	0.73	0.53	0.05
	IT	1	0.13	0.07	0.33	0.08	0.03
	TT	2	0.71	0.54	0.71	0.54	0.09

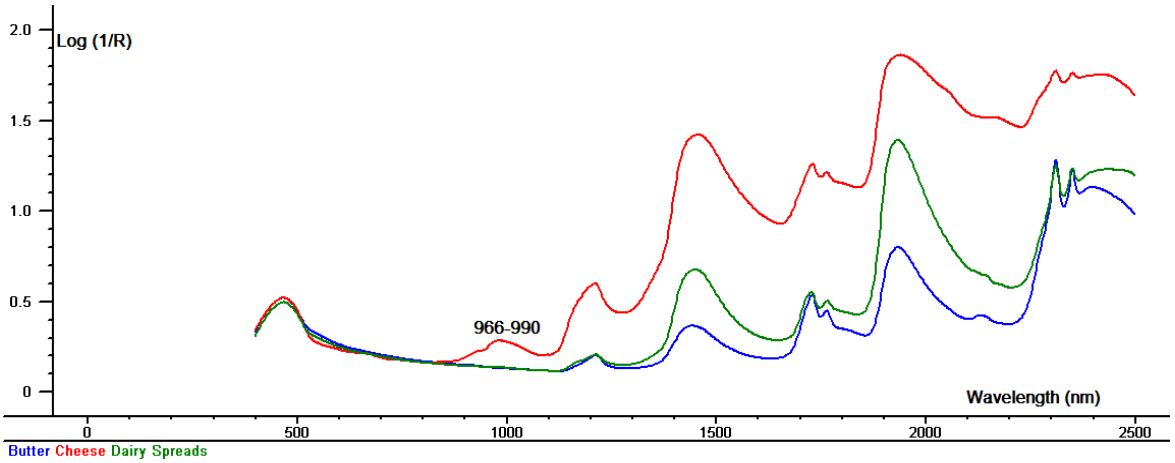
a)


Absorbance
0.15
0.10
0.05
0
1540
3380
Wave numbers (cm-1)
698.22974cm-1
1.24987e+03cm-1
1.80151e+03cm-1
2.35315e+03cm-1
2.90479e+03cm-1
3.45643e+03cm-1
Butter Cheese Dairy Spreads

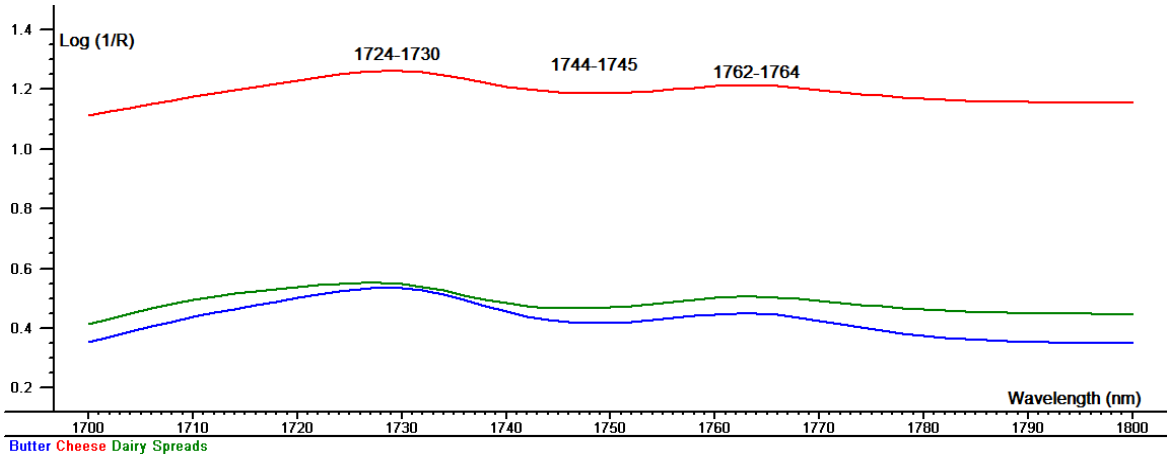
b)


Absorbance
0.15
0.10
0.05
0
966
1103-1117
1165
1242
1379
1415
1484
1540
1644-1654
1745
Wave numbers (cm-1)
900.75494cm-1
1.05120e+03cm-1
1.20165e+03cm-1
1.35210e+03cm-1
1.50254e+03cm-1
1.65299e+03cm-1
Butter Cheese Dairy Spreads

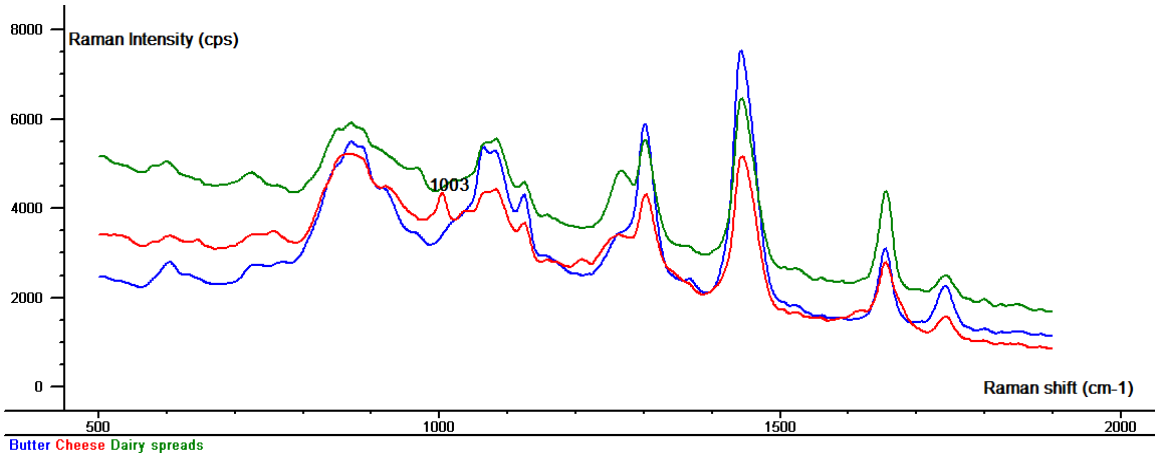
c)


Log (1/R)
2.0
1.5
1.0
0.5
0
966-990
Wavelength (nm)
0
500
1000
1500
2000
2500
Butter Cheese Dairy Spreads

d)



e)



f)

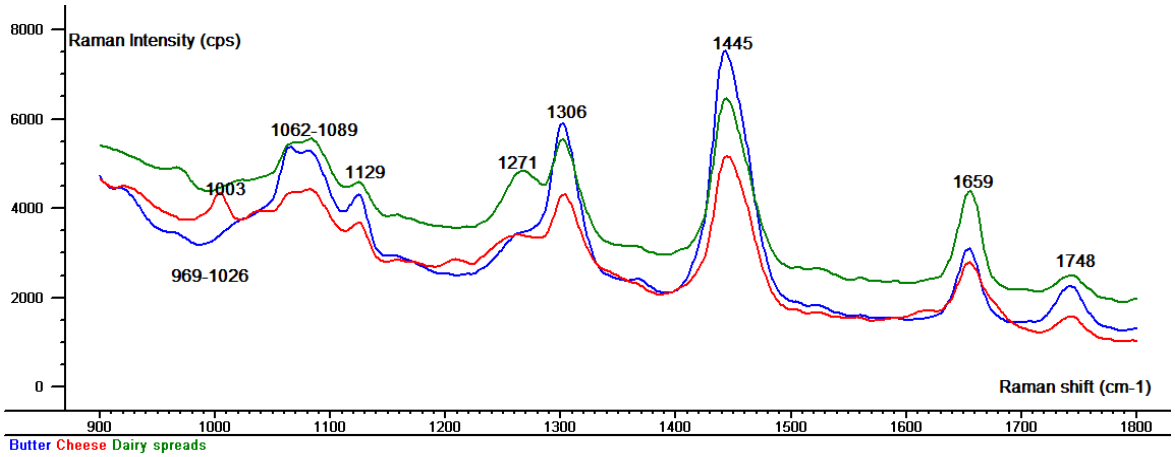


Figure1. Raw FT-MIR, NIR and Raman spectra of each type of dairy product a) FT-MIR (4000-698cm); b) FT-MIR (1800-900 cm^{-1}); c) NIR (400-2500 nm); d) NIR (1700-1800 nm); e) Raman(2000-500 cm^{-1});f) Raman (1798-900 cm^{-1}).

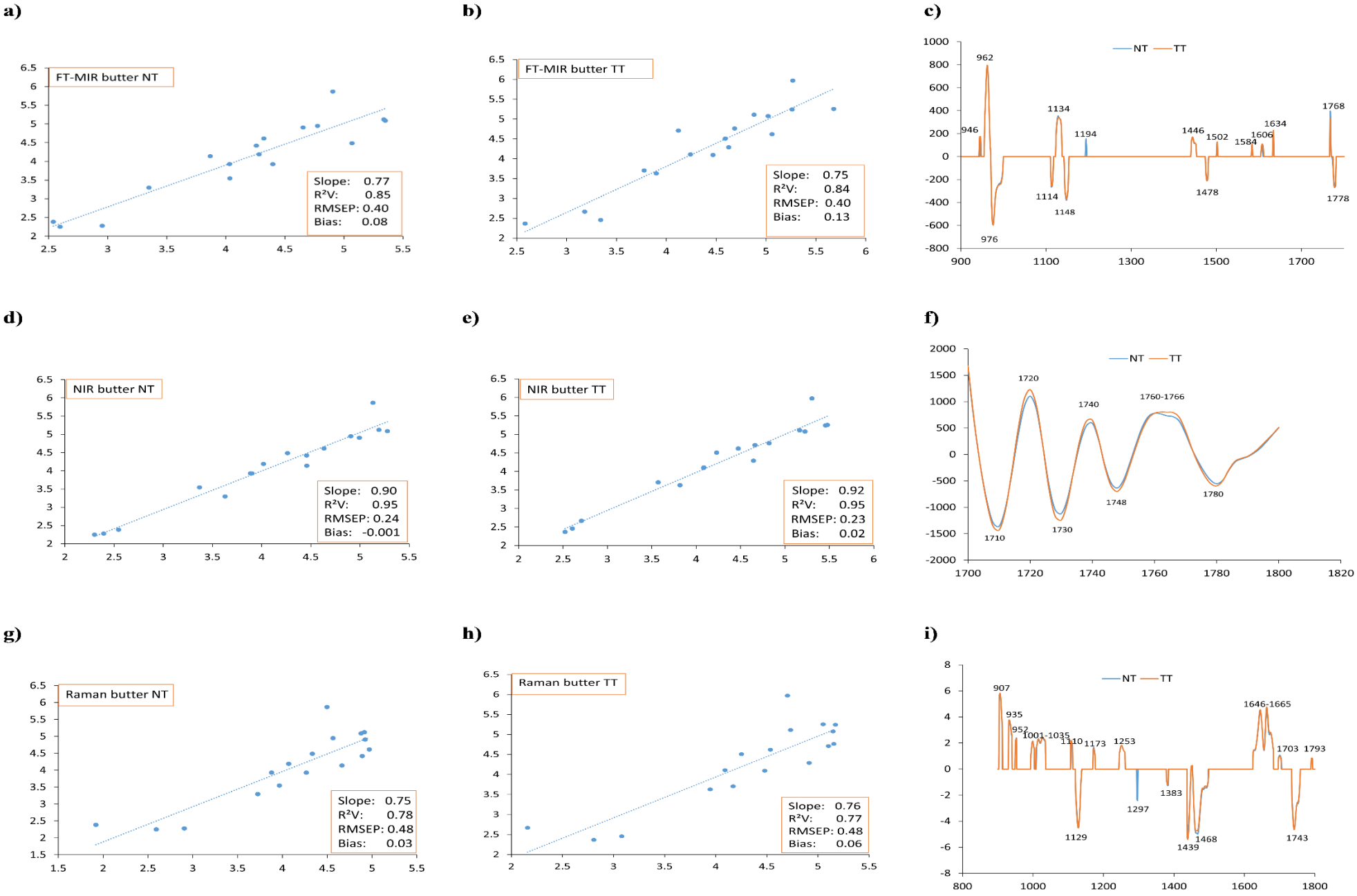
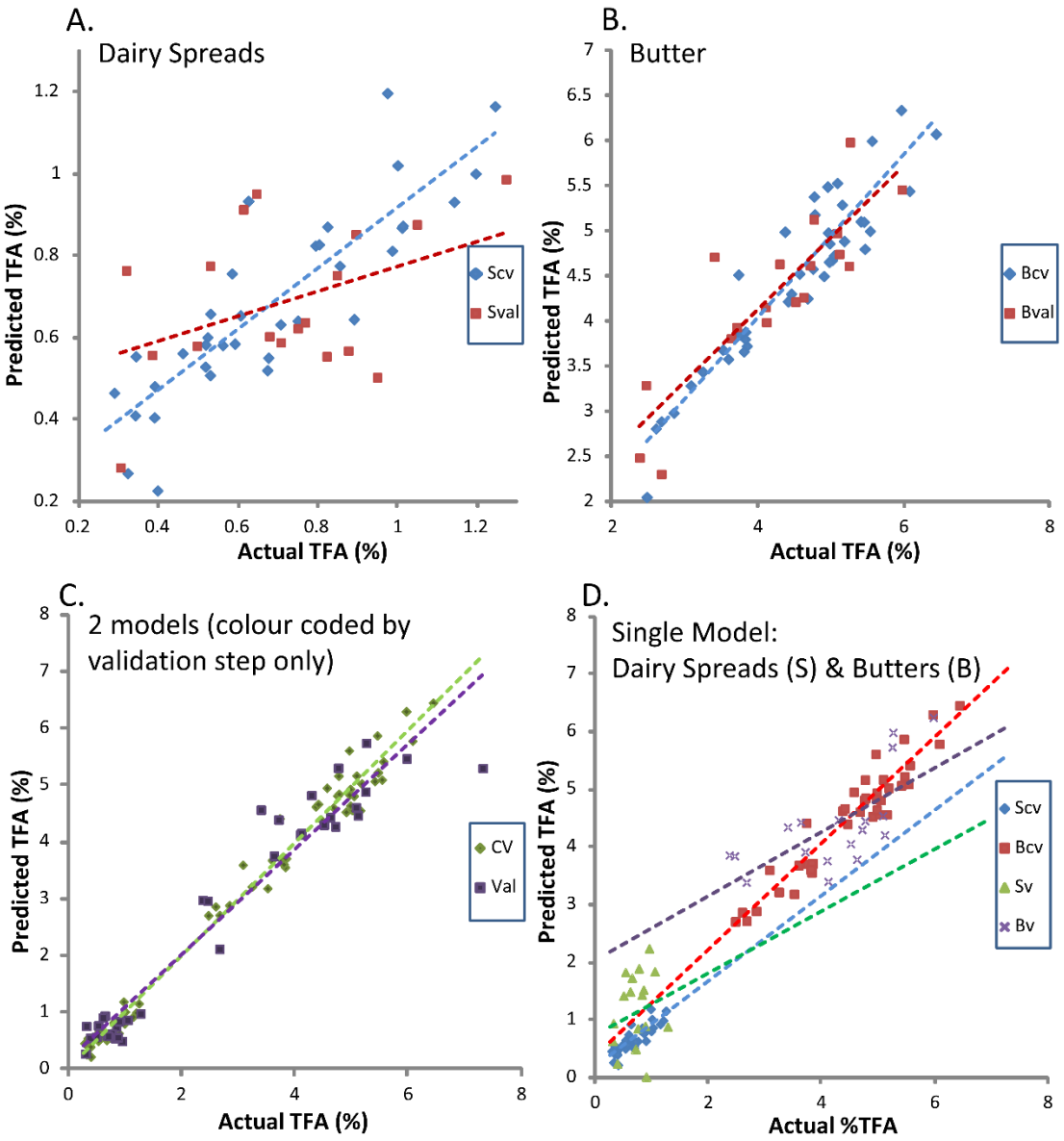

a)
FT-MIR butter NT
Slope: 0.77
R²V: 0.85
RMSEP: 0.40
Bias: 0.08
b)
FT-MIR butter TT
Slope: 0.75
R²V: 0.84
RMSEP: 0.40
Bias: 0.13
c)
NT
TT
962
946
976
1134
1114
1148
1194
1446
1478
1502
1584
1606
1634
1768
1778
d)
NIR butter NT
Slope: 0.90
R²V: 0.95
RMSEP: 0.24
Bias: -0.001
e)
NIR butter TT
Slope: 0.92
R²V: 0.95
RMSEP: 0.23
Bias: 0.02
f)
NT
TT
1720
1740
1760-1766
1710
1730
1748
1780
g)
Raman butter NT
Slope: 0.75
R²V: 0.78
RMSEP: 0.48
Bias: 0.03
h)
Raman butter TT
Slope: 0.76
R²V: 0.77
RMSEP: 0.48
Bias: 0.06
i)
NT
TT
907
935
952
1001-1035
1110
1173
1253
1129
1297
1383
1439
1468
1646-1665
1703
1793
1743

Figure 2. Linear regression plots of GC reference (x-axis) versus predicted (y-axis) values (% w/w of total fatty acids in FAME) for the validation set of NT and TT predictions in butter: a) & b) FT-MIR; d) & e) NIR; g) & h) Raman. Regression coefficients plots (wavelength vs. intensity (expressed in arbitrary units)), c) FT-MIR (900-1800 cm^{-1}); f) NIR (1700-1800 nm); i) Raman 900-1798 cm^{-1} for the prediction of NT and TT in butter.



	2 Models		Single Model	
	R^2_V	RMSEP	R^2_V	RMSEP
Spreads	0.178	0.201	0.036	0.695
Butter	0.80	0.32	0.473	0.618
Both	N/A	N/A	0.855	0.656

Figure 3. A and B: cross validation (blue) and independent test set regressions for total *trans* fatty acid content for the dairy spreads and butters. C: combined results for the two models plotted using a combined dataset, with best fit lines across both sample types. D: Single models based on individual datasets, colour coded according to the sample type with best fit trends for each sample type indicated.

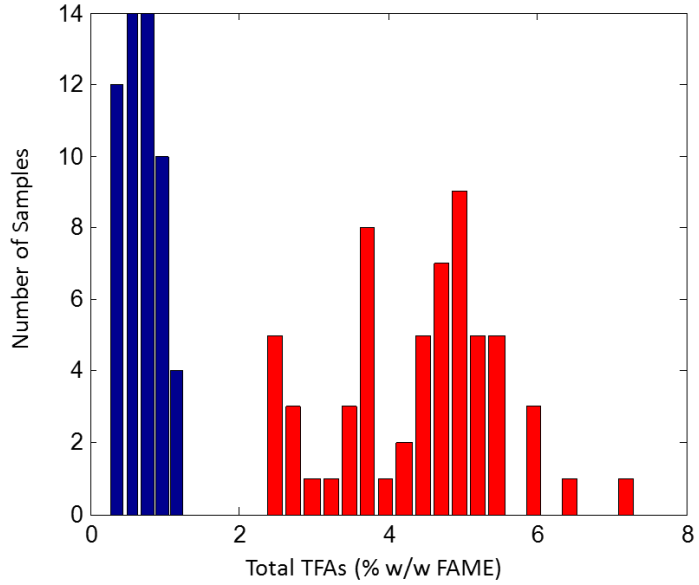


Figure 4. Distribution of total TFAs for the butter and dairy spread samples used in the study.

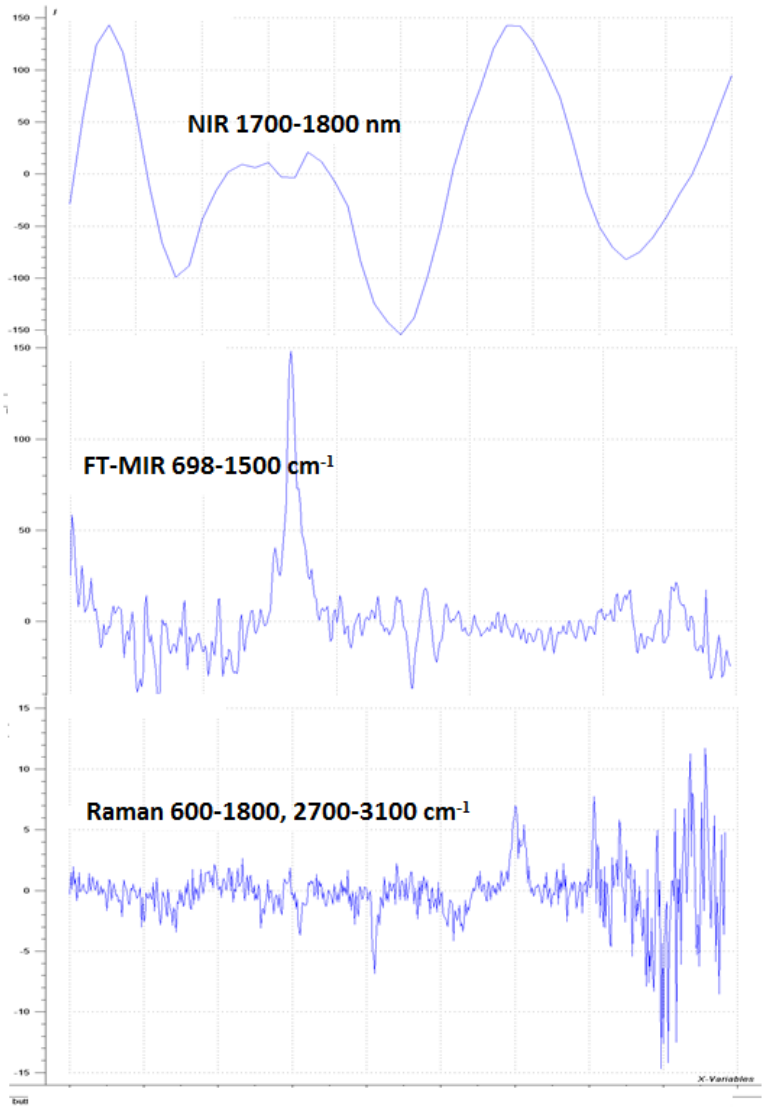


Figure 5. Regression coefficients without Martens uncertainty test for each of the three technologies used in the study.