TITLE: Exploration of microwave dielectric and near infrared spectroscopy with multivariate data analysis for fat content determination in ground beef

AUTHORS: Ming Zhao, Gerard Downey, Colm P. O'Donnell

NOTICE: This is the author's version of a work that was accepted for publication in *Food Control*. 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 *Food Control*, Available online 19 March 2016, http://dx.doi.org/10.1016/j.foodcont.2016.03.031

Exploration of microwave dielectric and near infrared spectroscopy with multivariate data analysis for fat content determination in ground beef

Ming Zhao[1, 2], Gerard Downey[1, 2], Colm P. O'Donnell [2*]

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

[2] School of Biosystems and Food Engineering, University College Dublin, Belfield, Dublin 4, Ireland

* Corresponding author. Tel: +353 1 7167201 E-mail address: colm.odonnell@ ucd.ie

Abstract

This study investigated using microwave dielectric and near infrared (NIR) spectroscopy for the determination of fat content in ground beef samples (n=69) in a designed experiment. Multivariate data analysis (principal component analysis (PCA) and partial least squares (PLS) regression modelling) was used to explore the potential of these spectroscopic techniques over selected multiple frequency or wavelength ranges. Chemical reference data for fat and water content in ground beef were obtained using a nuclear magnetic resonance-based SMART Trac analyser. Best performace of PLS prediction models for fat content revealed a coefficient of determination in prediction (R^2P) of 0.87 and a root mean square error of prediction (RMSEP) of 2.71% w/w for microwave spectroscopy; in a similar manner, R^2P of 0.99 and RMSEP of 0.71% w/w were obtained for NIR spectroscopy. The influence of water content on fat content prediction by microwave spectroscopy was investigated by comparing the prediction performance of PLS regression models developed using a single Y-variable (PLS1; fat or water content) and using both Y-variables (PLS2; fat and water contents).

Keywords

microwave dielectric spectroscopy, near infrared spectroscopy, multivariate data analysis, principal component analysis (PCA), partial least squares (PLS) regression modelling

1. Introduction

In recent years, spectroscopy has become a popular technology for food quality assessment due to its rapid data acquisition and ability to simultaneously predict multiple quality parameters with only minimal sample preparation. This technique has shown potential in quality control applications and determination of proximate composition of meat and meat products (Damez & Clerjon, 2013). Spectroscopy uses electromagnetic radiation to induce and measure vibrational or rotational energy transitions at molecular and atomic levels respectively. Based on the respective motions of quantum particles in the applied electromagnetic fields, this technique has been categorised into vibrational and rotational spectroscopic methods (Hollas, 2004).

Microwave-based dielectric spectroscopy is a rotational spectroscopic method which has been used to analyse qualitative characteristics of meat and other food materials since 1983 (Metaxas & Meredith, 1993). Theoretically, in a microwave-excited electromagnetic field, microwave radiation can easily penetrate food materials to induce particle polarisation in different components, giving rise to the characteristic dielectric properties of materials (Metaxas & Meredith, 1993). The dielectric property of a food is described as the complex permittivity which arises when it is subjected to an electromagnetic field in a microwave heating application. Complex permittivity (ε) is defined as:

$$\varepsilon = \varepsilon' - j\varepsilon'',$$

where ε' is the real part, called the dielectric constant, which represents the capacity of the material to store energy (Nelson & Datta, 2001); ε'' is the imaginary part referred to as the loss factor which is relevant to high frequency heating generated from electrical energy and includes the effects of conductivity-it also can be described as the ability of the material to convert energy into heat; *j* is the imaginary unit (j^2= -1). The ratio of ε''/ε' is known as the loss tangent (Nelson & Trabelsi, 2012). As individual parameters, ε' or ε'' can be used to represent dielectric properties of a material while both of them are frequency- and temperature- dependent (Metaxas & Meredith, 1993).

The potential of dielectric spectroscopy to detect content, state and activity of water together with detection of added water in meat products has previously been reported (Castro-Giráldez, Botella, Toldrá & Fito, 2010; Kent & Anderson, 1996; Clerjon, Daudin & Damez, 2003). Microwave measurements in reflection or transmission modes were also used to predict beef aging and fish freshness (Clerjon & Damez, 2007), meat quality in poultry and pork (Castro- Giráldez, Botella, Toldrá & Fito, 2010; Samuel, Trabelsi, Karnuah, Anthony & Aggrey, 2012) and a wide range of fat content (up to 50% w/w) in minced beef (Gunasekaran, Mallikarjunan, Eifert & Sumner, 2005; Ng et al., 2008). However, most studies have focused on profiling the dielectric properties of different materials over a specific frequency range (e.g. 0.3-3 GHz, 0-25 GHz, 0.2-12 GHz, 0.5-20 GHz) (Gunasekaran, Mallikarjunan, Eifert & Sumner, 2005; Ng et al., 2008; Kent & Anderson, 1996; Castro- Giráldez, Botella, Toldrá & Fito, 2010) under different temperature (Gunasekaran, Mallikarjunan, Eifert & Sumner, 2005; Ng et al., 2008) or during the aging process (Kent & Anderson, 1996; Castro-Giráldez, Botella, Toldrá & Fito, 2010). Some previous studies have explored chemometric analysis such as principal compoment analysis (PCA) and partial least squares regression (PLSR) for the prediction of moisture in poultry meat, scallops and pork but only based on each single selected frequency (Kent &

85 Anderson, 1996; Kent, Peymann, Gabriel & Knight, 2002). There have been no studies which
86 explored the use of different microwave frequency ranges for beef quality prediction using
87 multivariate data analysis.

88 As a mature vibrational spectroscopic technique, near infrared (NIR) spectroscopy combined
89 with multivariate data analysis has been widely used in meat analysis and implemented to
90 provide timely measurements in on-, in- and at-line applications (Weeranantanaphan,
91 Downey, Allen & Sun, 2011). NIR and visible light in transmission mode were used to
92 evaluate beef quality through detection and quantification of marbling (Ziadi, Maldague,
93 Saucier, Duchesne & Gosselin, 2012). NIR transmittance spectroscopy was also used to
94 measure the content of fat and other compositional parameters in fish (Wold & Isaksson,
95 1997; Nortvedt, Torrissen & Tuene, 1998; Solberg & Frendriksen, 2001; Xiccato, Trocino,
96 Tulli & Tibaldi, 2004). NIR reflectance spectroscopy has been used for identification of meat
97 species in homogenised meat muscle (Cozzolino & Murray, 2004) ; quality control on meat
98 trimmings (O'Farrell et al, 2010; Wold et al., 2011) and to detect adulterations in minced
99 beef and beefburgers (Morsy & Sun, 2013; Zhao, Downey & O'Donnell, 2013).

100 Commercial ground beef is derived from skeletal muscle, including adherent fatty tissues and
101 added beef fat trimmings. The European Community has issued a regulation for the quality
102 grades of typical ground beef (EU Regulation 853/2004) which are referred to as Class-I
103 'lean mince beef'and Class-II 'minced pure beef'. Based on the compositional criteria of fat
104 contnent, lean mince beef should contain< 7% w/w fat while minced pure beef should contain
105 <20% w/w fat (FSAI, 2013). The fat content of ground beef has always been a critical issue
106 with respect to accurate nutritional labelling, fraudulent industrial practice (e.g. addition of
107 extra fat) and consumer protection. Commercial ground beef contains both intramuscular and
108 additional fat. When beef is comminuted, all gross physical characteristics of the meat are
109 lost; therefore, timely processing analytical technology is required to assist quality control

and inspection. Comparing the mechanics of microwave and NIR spectroscopy, microwaves can obtain useful data from bulky and non-homogenised samples by significant penetration into the material while NIR radiation may only provide information about the surface of the material because of its shallow penetration depth. NIR is thus problematic for measurements of non-homogenised opaque solids. Among disadvantages of the microwave technique is the fact that the dielectric properties of moisture in samples may dominate or influence the dielectric spectral signals in certain frequency ranges (Kent & Anderson, 1996; Clerjon & Damez, 2009), as water has significantly higher permitivity than other organic compounds (Gunasekaran, Mallikarjunan , Eifert & Sumner, 2005).

The current study investigated the determination of fat content in ground beef using a laboratory-based dielectric measurement system and compared the results to those obtained on the same samples using a NIR spectroscopic method. Multivariate data analysis was utilised to develop and validate prediction models for ground beef fat and moisture contents. Performance of microwave and NIR prediction models was compared to evaluate the relative merits of each technique for use in ground beef quality control. This study is innovative and novel for the following reasons: 1) this study involved a greater number of ground beef samples with more variables than previous research; 2) the determination of fat content in ground beef samples has been explored using multivariate data analysis using different dielectric frequency ranges, and 3) a chemometric method using PLS-1 and PLS-2 modelling has been introduced to microwave spectroscopy to examine the potential influence of water content on the determination of fat content in ground beef.

2. Methods and Materials

134 *2.1 Sample preparation*

135 Ground beef samples (n=69) were produced in the abbatoir of Teagasc Food Research Centre
136 (Ashtown, Dublin 15, Ireland). These samples were produced to include both a higher and a
137 lower quality level in beef. In the higher quality group, round steak beef cuts were trimmed to
138 remove visible fat; beef fat trimmings were added to aliquots of lean beef to produce samples
139 (0.5 kg each) with 0%, 3%, 5% and 7% w/w of added fat content. In the lower quality group,
140 plate or flank cuts were trimmed of visible fat but connective tissue and tendons were left.
141 Lower quality samples (0.5 kg each) were produced using these trimmed cuts with added
142 beef fat trimmings at 5% to 30% w/w in 2.5% w/w increments. Samples at each added fat
143 level were produced on three separate occasions using meat sourced from three different
144 butcher shops. In all, 36 higher quality samples and 33 lower quality samples (primary
145 samples) were produced using raw materials from 9 local butchers. During preparation, each
146 sample was ground twice using a Mainca meat mincer (PM70/12; Cheshire, UK) followed by
147 homogenisation for 1 minute using the Robot coupe R301 ultra (Vincenes, France).
148 Homogenised samples were transferred into lidded plastic containers and stored at +3°C.
149 Before measurement, samples were allowed to equilibrate to ambient laboratory temperature
150 (20 ± 2°C).

151

152 *2.2 Sample measurements*

153 *2.2.1 Measurements dielectric*

154 A time-domain reflectometry dielectric system was used in this study. Dielectric spectra were
155 derived from the reflected signals in an electro-magnetic field. Dielectric properties (i.e.
156 dielectric constant (ε'), dielectric loss factor (ε'') and loss tangent (ε''/ε')) were recorded over
157 the 200 MHz to 20 GHz frequency range using an Agilent 85070E high temperature coaxial
158 probe kit connected to an impedance analyser (E4991A; Agilent Technologies, Santa Clara,

159 USA). This probe featured a hermetic glass-to-stainless steel 19 mm diameter seal and a 3.5
160 mm aperture; this configuration requires homogeneous samples with a flat, gas-free surface
161 to ensure contact at the probe face during measurement and a certain minimum sample
162 thickness, i.e., $>\sqrt{|\varepsilon|}$ mm (Keysight technologies, 2014). A probe stand was used to fix the
163 probe and the connected coaxial cable in a constant position as even subtle changes in
164 instrument configuration may affect calibration; the sample, contained in a holder, was raised
165 to reach the probe (Fig. 1). A daily calibration check was performed using deionised distilled
166 water and a short circuit flange before collection of any measurement data. In this study,
167 homogenised beef samples were manually compressed into disposible plastic containers
168 (volume ~ 100 ml) to minimise air bubbles and form a flat upper surface. Two sub-samples
169 were taken from each primary minced beef sample and three measurements on different
170 areas of the surface were taken for each sub-sample. These replicate measurements were
171 averaged prior to data analysis. The permittivity of each sample was derived from the
172 reflection coefficient as a function of the frequency measured by the impedance analyser and
173 converted to dielectric behaviour of ε', ε'' and $\varepsilon''/\varepsilon$ at the frequency range using numerical
174 models embedded in the manufacturer's software (Agilent technologies, 2006).
175
176 *2.2.2 Near infrared spectroscopic measurements*
177 Reflectance spectra [log (1/*R*)] of each sample were recorded using a NIRSystems 6500
178 instrument (NIRSystems Inc., Laurel, MD, USA) over the wavelength range 400–2498 nm (2
179 nm intervals) and using a 16–32 reference-sample scan sequence. Samples were scanned in
180 random order at ambient temperature (~20 °C). Two sub-samples of each primary sample
181 was taken for the measurements; each sub-sample was scanned three times with rotation of
182 the sample cup through ~120° between the sequential scans.The mean of these repetitions
183 was used in later chemometric operations. Instrument control, spectral collection and file

184 export was performed using WINISI software (v.1.04; Intrasoft International, State College,
185 PA, USA).
186
187 *2.2.3 Determination of fat content*
188 Fat content of each sample was determined using a bench-top nuclear magnetic resonance
189 (NMR) instrument (SMART Trac Fat Analyser; CEM Corporation USA). Approximately 3.5
190 g of each sample was initially dried using the SMART Trac microwave drying oven and the
191 moisture content of each sample was recorded in this step. Afterwards, fat content was
192 estimated by NMR radio frequency induced signals of hydrogen on the basis of proton decay
193 in the fatty acids (AOAC, 2000). Based on the formula of each sample, fat content was
194 predicted using the supplied NMR Smart trac calibrations for high (30-70% w/w) and low
195 (0-20% w/w) fat contents. In this study, based on the designed approximate fat content levels
196 (0-30% w/w) in the sample production, samples with less than 20% of added fat were tested
197 using the low fat calibration and others with 20-30% of added fat were analysed using both
198 the high and low fat methods to get the average value of the predicted results. Each sample
199 was tested in triplicate by the selected methods and the mean was used as the chemical
200 reference value of each sample for the development of spectroscopic prediction models.
201
202 *2.3 Multivariate data analysis*
203 Dielectric constant (ε'), loss factor (ε'') and loss tangent (ε''/ε') data were saved in the
204 manufacturer's file (.prn) format and converted into Excel (Microsoft Office 2010) files. NIR
205 spectra were exported from WINISI software as JCAMP.DX files. All data files were
206 imported into The Unscrambler software (v.9.7; Camo, Trondheim, Norway) for
207 chemometric operations. Data were pre-treated using normalisation on unit vectors (nor.u.v.),
208 multiplicative scatter correction (MSC), standard normal variate (SNV) transformation and

209 Savitzky Golay derivatisation to remove baseline shifts, slope changes, scatter effects and
210 reveal greater structure in the spectral data. Principal component analysis (PCA) was
211 performed initially on the raw spectral data of the whole sample set using leave-one-out
212 cross-validation to investigate the presence of sample clustering.
213 Partial least squares regression (PLSR) models were developed on spectral data (X-values)
214 and chemical reference data (Y-values) for the prediction of fat content in ground beef. In
215 dielectric spectroscopy, dielectric constant (ε'), loss factor (ε'') and loss tangent (ε''/ε') over
216 a number of frequency ranges (0.6-20 GHz, 0.6-3 GHz, 0.6-9.7 GHz, 8.1-12.1 GHz, 2.6-20
217 GHz, 10.1-20 GHz, 12.9-20 GHz and 16.1-20 GHz) were used for the chemometric
218 operations. NIR spectra over selected wavelength ranges, i.e. 780-2500 nm, 1400-2500 nm
219 and 900-1800 nm were also explored in this study. Spectral range selections were based on
220 the choices of previous studies and by examination of the intensities of regression
221 coefficients. The Martens uncertainty test (Martens & Martens, 2000) was applied to develop
222 PLSR models using a reduced number of spectral variables in an effort to improve model
223 stability. A PLS prediction model with two Y-variables (PLS-2 with Y_1 (fat content) and Y_2
224 (moisture content)) was developed over the entire frequency range (0.6-20 GHz) to explore
225 the influence of water content on fat content prediction. In this case, fat and moisture content
226 information were simutaneously modelled in the PLS regression.
227 All ground beef samples were numbered from 1 to 69 in order of production. Samples with
228 even numbers were chosen arbitrarily as the calibration data set while the remainder formed a
229 separate validation set. The same calibration and validation sample sets were used in the
230 chemometric operations of both spectroscopic methods.
231 PLSR model performances were evaluated on the basis of the magnitude of the cross-
232 validation coefficients of determination, root mean square errors in calibration (i.e. R^2CV,

RMSECV) and their validation counterparts (i.e. R^2P, RMSEP). Bias values and slopes of the fitted regression lines were also calculated. R^2, RMSE and bias were calculated as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2} \quad (1)$$

$$\text{RMSE} = \left(\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}\right)^{0.5} \quad (2)$$

$$\text{Bias} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)}{n} \quad (3)$$

Where $\hat{y}_i$ and y_i are the predicted and reference values of one sample (i); $\bar{y}$ is the mean of the reference values of all the samples. The value of R^2 is expected to be close to 1 while values of RMSE and bias are expected to be close to 0.

PLS regression modelling finds a set of latent vectors by projecting response variables (Y) and independent variables (X) onto a new subspace while maximising the collinearity between the orthogonal scores of X and Y (Wold, Sjöström & Eriksson, 2001). PLS-1 and PLS-2 are two different algorithms, which are used to handle single and multiple response variables (Y) with the same independent variable (X) respectively (Wold, Martens & Wold, 1983). PLS-1 and PLS-2 are expressed as simple formulae as follows:

$$Y_1 = X B_1 + E_1 = X \hat{W}_a S_a + E_1 = X (P_a^T)^{-1} S_a + E_1 \quad (4)\text{-------- PLS-1}$$

$$Y_2 = X B_2 + E_2 = X \hat{W}_b \hat{S}_b + E_2 = X (P_b^T)^{-1} (Q_b^T)^{-1} + E_2 \quad (5)\text{-------- PLS-2}$$

$$\hat{W}_n = W_n (P_n^T W_n)^{-1} \quad (6)$$

$$\hat{S}_b = S_b (Q_b^T S_b)^{-1} \quad (7)$$

Where respectively, B_1 and B_2 are the regression coefficients of PLS-1 and PLS-2; E_1 and E_2 are squares error matrices of PLS-1 and PLS-2; $\hat{W}_a$ and $\hat{W}_b$ are the weights from spectral data using PLS-1 and PLS-2; a and b are the numbers (n) of latent variables for PLS-1 and PLS-2.

254 $\hat{S}_b$ is the weights from the residuals of multiple chemical compounds using PLS-2. P_a and S_a
255 are the loadings of X and Y_1 for PLS-1; P_b and Q_b are the loadings of X and Y_2 for PLS-2 ; T
256 represents transpose of a matrix.
257 In this study, the single response variable (Y) of PLS-1 was the reference value of fat content
258 in ground beef; the multiple response variables (Y_1 and Y_2) of PLS-2 were the reference
259 values of fat and water content respectively; the independent variable (X) was a matrix of
260 spectral data of all samples. If the prediction value of fat content was affected by the water
261 content, then the important statistical parameters (e.g. R^2CV, RMSECV, R^2P and RMSEP) of
262 the PLS-1 model would show significant difference from those of the PLS-2 model. As in the
263 current study, the PLS-2 model involved the two loading vectors related to the two Y
264 variables in the sample matrix, the projection of their combined score matrix onto a space
265 would be different from that of each individual score matrix, except in the case that the two Y
266 variables were strongly inter-correlated and shared the same loading vector.
267
268
269 # 3 Results and Discussions
270 *3.2 Fat and moisture content of ground beef samples*
271 Fig. 2a shows fat and water content values (% w/w; y-axis) of the ground beef samples
272 (n=69; x-axis) estimated using the SMART Trac Fat Analyser. Sample fat content does not
273 match exactly the planned value from the experimental design due to the presence of
274 intramuscular fat and moisture contained in beef fat trimmings; however, fat and water
275 content values recorded by the Smart Trac instrument do include intramuscular (~ 1-3% w/w)
276 and added moisture from beef fat trimmings. Measured fat and water contents of each sample
277 are shown in Fig. 2a; this data approximately follows the experimental design for both
278 minced pure beef and lean beef mince samples. Fig. 2b shows the strong correlation (R^2 ~

279 0.97)) between fat and water content in these samples. In this study, fat and water content of
280 all the tested samples ranged from 0.8 to 30.6% w/w and from 56.2 to 74.2% w/w
281 respectively. The standard deviation of the measured fat or water content in each sample was
282 derived from two measurement repetitions. Over the standard deviations of all the samples,
283 the measurement repeatability (repeatability of standard deviation) for fat content was
284 estimated to be 0.46% w/w; the estimated repeatablility for water content was 0.37% w/w.
285
286 *3.3 Dielectric and NIR spectra with PCA analysis*
287 Dielectric spectra of the complex dielectric constant (ε) were separately expressed as the
288 spectra of dielectric constant (ε'), loss factor (ε'') and loss tangent (ε''/ε') over the 0.6 to 20
289 GHz frequency range (Fig. 3a, b and c). In this case, dielectric spectra (Fig. 3a, b, c) exhibited
290 two different patterns; those in Fig. 3a displayed a vertical offset which varied from generally
291 larger to smaller over the measured range whereas those in Fig. 3b and c displayed a marked
292 initial decrease in spectral intensity in the range 0.6 to approximately 2.0 GHz followed by an
293 increase over the rest of the spectral range. The spectra in Fig. 3b and c exhibited the same
294 behaviour whereas in Fig. 3a, it appears that, while the majority of spectra exhibit a decline in
295 intensity over the frequency range, some revealed an increase from a minimum at 0.6 GHz to
296 a broad maximum centred around 11-11.5 GHz. On the whole, these spectra are relatively
297 feature-less with the most striking feature being the vertical offsets. Among reasons which
298 may be advanced to explain these phenomena may be the distance of the measurement probe
299 head to the sample surface as the dielectric system used in this study was sensitive to a couple
300 of millimeter difference between sample heights (Agilent technologies, 2006) and
301 interference due to the unavoidable presence of air sacs under the sample surface that caused
302 energy loss when microwave passed through these empty air sacs.

NIR reflectance spectra over the 400-2498nm wavelength range are shown in Fig. 5a. Generally all the spectra showed absorbance peaks at almost the same wavelength locations. One peak at 970nm may be attributed to the second overtone of O-H absorbance and corresponds to water content. Peaks at 1940nm (a combination of O-H stretch and deformation) and 1450 nm (the first overtone of O-H stretch vibration) have also been ascribed to water. Other peaks at 1200 nm (the second overtone of –CH stretch), 1728 and 1765 nm (the first overtone of –CH stretch) and 2310-2345 nm (combined –CH stretch and deformation band) have been related to fat moieties (Osborne & Fearn, 1988).

PCA analysis was used to identify the major sources of variance based in these spectral data and to highlight any unusual or outlying samples. In Fig. 4a, b, c and Fig. 5b, the first two PC factors (PCs 1 & 2) explained 98-99% of the spectral variance and only one major cluster was defined in this 2-dimensional space. The Hotelling T^2 ellipse was used to detect potentially outlying samples on the basis of sample Mahalanobis distance from the PCA model using a 95% confidence limit. The vast majority of all samples was contained in the ellipse in each case with only a very small number of samples (i.e. b3, b9 in Fig. 4a; b3 in Fig. 4b; b3, b9, b10 in Fig. 4c and b2 in Fig. 5a) lying a small distance outside; on this basis, no samples were excluded from the subsequent PLSR data analysis.

3.4 Fat content prediction models using dielectric and NIR spectral data

PLSR models with one y-variable (fat content; PLS1) were developed using full cross-validation on spectral data (dielectric ε', ε'' , ε''/ε' and NIR spectra) of the calibration sample group (n=35) and subsequently validated on the remaining samples (n=34). Tables 1 and 2 show summary statistics of model performance; all the above-mentioned frequency ranges and data pre-treatment methods were investigated on these datasets but only the models with good performance are displayed here.

328 Table 1 shows that the most accurate models based on dielectric spectra involved both raw
329 and normalisation on unit vector pre-treated data, normally involving fewer than 6 PLS
330 loadings. In general, it may be stated that prediction models developed using ε' spectral data
331 showed slightly better performance than those developed using ε'' or ε''/ε' data. Values of
332 R^2CV and R^2P ranged from 0.73 - 0.81 and 0.68-0.87 respectively; RMSEP is in the range of
333 2.71-4.28 % w/w with bias values of between -0.47 and -0.86 in ε' spectra based models. The
334 most accurate dielectric model was developed using normalised ε'data over the 12.9-20
335 GHz frequency range with the slope of the line-of-best-fit equal to 0.88, R^2CV of 0.81,
336 RMSECV of 3.71, R^2P of 0.87 and RMSEP of 2.71. (Fig. 6a). In the case of the parameters
337 ε'' and ε''/ε', the frequency ranges which produced the most accurate calibrations were 0.6-
338 20 GHz and 0.6-9.7 GHz respectively using raw spectral data. R^2CV and R^2P values of 0.82
339 and 0.80 respectively together with RMSECV and RMSEP values of 3.37 and 3.61 were
340 obtained using ε'' measurement; in a similar manner, the most accurate ε''/ε' model had an
341 R^2CV of 0.80, RMSECV of 3.81, R^2P of 0.77 and RMSEP of 3.66. Fig. 6b shows the
342 regression coefficient intensities of the most accurate model for fat prediction using the ε'
343 parameter. High intensity values appeared at 13.66, 14.46, 16.04 and 18.81 GHz. At the
344 present time, the authors are unaware of any studies in which the relationship between
345 chemical components of a mixture and absorption at specific frequency values has been
346 clearly described; therefore, the chemical basis for the features in this or any other regression
347 model developed in this work may not be determined as yet.
348 As the dominant constituent, water in meat may interfere with prediction of fat content over a
349 certain frequency range. An early study reported that water content of meat had been
350 correlated with dielectric constant (ε') over the low frequency range (0.2-12 GHz) and
351 especially at a frequency close to 2 GHz (Kent and Anderson 1996). In the current study, a
352 PLS prediction model with two y-variables (PLS-2 with fat content and moisture content)

was developed over the selected frequency range (0.6-20 GHz) to explore the influence of water content on fat content prediction. Three PLS loadings were optimal in the developed model and the relevant regression coefficient plots are shown in Fig. 7a, b and c. Those modelling fat content (Fig. 7a) displayed a trend of increasing intensity over the frequency range whereas those for water content showed the inverse behaviour. For the second and third PLS loadings, regression coefficient intensities for fat content decreased over 0.6-9.7 GHz and increased over 9.7-20 GHz while those for water content again exhibited inverse behaviour (Fig. 7b, c). Overall, highest intensity of regression coefficients for fat content prediction appeared at a higher frequency range while that for moisture content prediction appeared at a lower frequency range. To investigate the possibility that water and fat interfered with each other, a PLS-2 regression was performed. The results of both the PLS-1 and PLS-2 analyses are summarised in Table 3 where it can be seen that almost identical accuracies were achieved by both strategies. Values of R^2CV, RMSECV, R^2V and RMSEP from PLS-1 models are almost as the same as the values from the PLS-2 model; differences of RMSE values between PLS-1 and PLS-2 were less than 0.06. On the other side, in Fig. 2, the reference values of fat and water content were strongly correlated to each other with a correlation coefficient determination (R^2) of 0.97. It assumed that the two score matrices of two Y variables (i.e. fat and water) shared a similar loading vector. Therefore, in the current study, water content in ground beef was shown not to interfere with the fat content prediction using dielectric spectroscopy.

NIR models were developed and validated using the same calibration and validation sample sets as the dielectric models; in general, this technique revealed excellent performance in the prediction of fat content in ground beef. Table 2 reveals the performance of NIR models after application of the Martens uncertainty test over the full NIR range (780-2500 nm), the range mainly related to fat content (1400-2500 nm) and the range of water, protein and fat (900-

1800nm). These models were developed using raw data and data pre-treated with SNV, MSC and normalisation and used 2-4 PLS loadings. Overall, models in Table 2 show high values of slope (0.97-1), R^2CV (0.97-0.99), R^2P (0.96-0.99) and low values of both RMSEP (0.67-1.46) and bias (- 0.22 – 0.15). The best performing model was based on raw spectral data over the 1400-2500nm wavelength range using 4 PLS loadings with both R^2CV and R^2P of 0.99, RMSEV of 0.84 , RMSEP of 0.71 and a slope value of 1 (Fig. 6c). Regression coefficients associated with this model are shown in Fig. 6d. Only a small number (n=176) of wavelengths were retained after application of the Martens uncertainty test. Two important peak areas of regression coefficient intensity can be seen at 1726 nm and 2310 nm which are separately related to the first overtone of –CH stretch, the combined –CH stretch and deformation band of aliphatic chains in fatty acids (Osborne & Fearn, 1988). Other obvious intensity changes happened over 1554-1660 nm and 2080-2224nm which contain complementary information concerning N-H combination bands from amides or proteins and O-H band of water, C-H/C=O stretching combinations of lipids (Workman & Weyer, 2008).

3.5 Sample error discussion

In the current study, as the SMART Trac System (a secondary analysis method) was used to determine the chemical reference values of fat content. Repeatability errors from the SMART Trac analyser were calculated to be 0.46% w/w for fat; by definition, this error will transfer the estimation of RMSEP by microwave or NIR spectroscopy. Additionally, for the dielectric spectroscopic system, some specific measurement errors related to the predicted difficulties of preparing a sample with a perfect flat surface that no further than 2.5 μm from the flat probe face (Agilent technologies, 2006) and also without empty air sacs contained in the sample. All these errors mitigated against achievement of the minimum prediction error for

both spectroscopic techniques; however, the accuracies reported in this paper represent practically-achievable values.

4 Conclusions

Prediction models for fat content in ground beef samples prepared in this designed experiment were derived by multivariate data analysis of microwave and NIR spectra. In the case of microwave spectroscopy, the best predictive performance used dielectric spectral data (ε') was over a 12.9-20 GHz range; with an R^2 value of 0.87 and RMSEP equal to 2.71% w/w. The authors suggest that time-domain reflectometry microwave spectroscopy has potential for quality control of ground beef products at a screening level. In the case of NIR spectroscopy on the same ground beef samples, the best predictive model produced an R^2P value of 0.99 and RMSEP equal to 0.71 % w/w using spectral data collected over the 1400 to 2500 nm range and after application of the Martens uncertainty test. Multivariate data analysis has rarely been used in previous studies of microwave spectroscopy; in this study, dielectric spectral information over different frequency ranges has been explored for fat content prediction in ground beef using PLS-1 and PLS-2 modelling. It has been found that water content in ground beef samples is highly inter-correlated with fat content but does not appear to interfere with fat content prediction, at least when the entire frequency range available was used (0.6-20GHz). Because homogenised ground beef samples were used in this work, the penetration depth advantage of microwave spectrocopy over NIR spectroscopy was not apparent here. A further study will investigate this advantage at an industrial processing level using unhomogenised ground beef for quality control purposes.

Acknowledgements

The authors acknowledge funding for this work from the Irish Department of Agriculture, Food and the Marine under the Food Institutional Research Measure (FIRM) programme over 2013-2017. The authors also would like to thank Prof. James Lyng for providing the dielectric equipment and his valuable suggestions on this work. Views expressed in this paper are those of the authors solely. Ming Zhao is a Teagasc Walsh Fellow.

References

Agilent Technologies. (2006). Agilent basics of measuring the dielectric properties of materials (Application note). Available from: http://www3.imperial.ac.uk/pls/portallive/docs/1/11949698.PDF. Accessed 2015 June 27.

AOAC. (2000). Official methods of analysis, method no. 985.14. AOAC INTERNATIONAL, Gaithersburg, MD: Association of Official Analytical Chemists.

Castro-Giráldez M., Botella P., Toldrá F. & Fito P. (2010). Low-frequency dielectric spectrum to determine pork meat quality. *Innovative Food Science and Emerging Technologies. 11,* 376-386.

Clerjon S., Daudin J. D., & Damez, J. L. (2003). Water activity and dielectric properties of gels in the frequency range 200 MHz-6 GHz. Food Chemistry, 82, 87-97.

Clerjon S. & Damez J. L. (2007). Microwave sensing for meat and fish structure evaluation. *Measurement Science and Technology. 18,* 1038-1045.

Clerjon S. & Damez J. L. (2009). Microwave sensing for an objective evaluation of meat ageing. *Journal of Food Engineering. 94,* 379-389.

450 Cozzolino D. & Murray I. (2004). Identification of animal meat muscles by visible and near
451 infrared reflectance spectroscopy. *Food Science Technology-LEB. 37***,** 447-452.

452 Damez J. L. & Clerjon S. (2013). Quantifying and predicting meat and meat products quality
453 attributes using electromagnetic waves: An overview. *Meat Science. 95,* 879-896.

454 FSAI (Food Safety Authority of Ireland). (2013). Requirements for raw materials-minced
455 meat. Available from:
456 www.fsai.ie/legislation/food_legislation/minced_meat/raw_materials_minced_meat.html.
457 Accessed 2015 June 27.

458 Gunasekaran N., Mallikarjunan P., Eifert J. & Sumner S. (2005). Effect of fat content and
459 temperature on dielectric properties of ground beef. *Transactions of the ASAE. 48***,** 673-680.

460 Hollas J. M. (2004). Rotational spectroscopy. Vibrational spectroscopy. In: *Modern*
461 *spectroscopy* (4th ed.) (pp. 103-196), Chichester, West Sussex, England: John Wiley & Sons.

462 Keysight Technologies. (2014). Keysight 85070E Dielectric Probe Kit 200 MHz to 50 GHz.
463 Available from: www.keysight.com. Accessed 2015 June 27.

464 Kent M. & Anderson D. (1996). Dielectric studies of added water in poultry meat and
465 scallops. *Journal of Food Engineering. 28,* 239-259.

466 Kent M., Peymann A., Gabriel C. & Knight A. (2002). Determination of added water in pork
467 products using microwave dielectric spectroscopy. *Food Control. 13,* 143-149.

468 Martens H. & Martens M. (2000). Modified jack-knife estimation of parameter uncertainty in
469 bilinear modelling by partial least squares regression (PLSR). *Food Quality Preference. 11,*
470 5–16.

471 Metaxas A. C. & Meredith R. J. (1993). Dielectric properties. In: *Industrial microwave*
472 *heating* (pp. 5-66). IEE power engineering series 4. London: Peter Peregrinus.

473 Morsy N. & Sun D.-W. (2013). Robust linear and non-linearmodels of NIR spectroscopy for
474 detection and quantification of adulterants in fresh and frozen – thawed minced beef. *Meat*
475 *Science. 93***,** 292–302.

476 Nelson S. O. & Datta A. K. (2001). Dielectric properties of food materials and electric field
477 interactions. In: *Handbook of microwave technology for food applications* (pp. 69-114), New
478 York: Marcel Dekker.

479 Nelson S. O. & Trabelsi S. (2012). Factors influencing the dielectric properties of agricultural
480 and food products. *Journal of Microwave Power and Electromagnetic Energy. 78***,** 93-107.

481 Ng S. K., Ainsworth P., Plunkett A., Haigh A. D., Gibson A. A. P., Parkinson G. & Jacobs G.
482 (2008). Determination of added fat in meat paste using microwave and millimetre wave
483 techniques. *Meat Science. 79,* 748-756.

484 Nortvedt R., Torrissen O. J. & Tuene S. (1998). Application of near-infrared transmittance
485 spectroscopy in the determination of fat, protein and dry matter in Atlantic halibut fillet.
486 *Chemometrics and intelligent laboratory. 42,* 199-207.

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

489 Samuel D., Trabelsi S., Karnuah A. B., Anthony N. B. & Aggrey S. E. (2012). The use of
490 dielectric spectroscopy as a tool for predicting meat quality in poultry. *International Journal*
491 *of Poultry Science. 11*, 551-555.

492 Solberg C. & Frendriksen G. (2001). Analysis of fat and dry matter in capelin by near
493 infrared transmission spectroscopy. *Journal of Near Infrared Spectroscopy. 9***,** 221-228.

494 Weeranantanaphan J., Downey G., Allen P. & Sun D.-W. (2011). A review of near infrared
495 spectroscopy in muscle food analysis: 2005–2010. *Journal of Near Infrared Spectroscopy*.
496 *19*, 61–104.

497 Wold J. P. & Isaksson T. (1997). Non-destructive determination of fat and moisture in whole
498 Atlantic salmon by near-infrared diffuse spectroscopy. *Journal of Food Science. 62*, 734-736.

499 Wold S., Martens H. & Wold H. (1983). The multivariate calibration problem in chemistry
500 solved by the PLS method. In: Proc. Conf. Matrix Pencils, Ruhe A., Kågström B., (Eds.).
501 *Lecture Notes in Mathematics* (pp. 286-293). Heidelberg, Gremany: Springer Verlag.

502 Wold S., Sjöström M. & Eriksson L. (2001). PLS-regression: A basic tool of chemometrics.
503 *Chemometrics and intelligent laboratory*. *58,* 109-130.

504 Workman J. & Weyer J. L. (2008). Appendix 4a: spectra-structure correlations for near
505 infrared. In: *Practical Guide to Interpretive Near-Infrared Spectroscopy* (pp. 63–65). Boca
506 Raton, FL, US: Taylor & Francis Group.

507 Xiccato G., Trocino A., Tulli F. & Tibaldi E. (2004). Prediction of chemical composition and
508 origin identification of European sea bass (Dicentrarchus labrax L.) by near infrared
509 reflectance spectroscopy (NIRS). *Food Chemistry*. *86*, 275-281.

510 Zhao M., O'Donnell C. P. & Downey G. (2013). Detection of offal adulteration in
511 beefburgers using near infrared reflectance spectroscopy and multivariate modelling. *Journal*
512 *of Near Infrared Spectroscopy. 21***,** 237−248.

513 Ziadi A., Maldague X., Saucier L., Duchesne C. & Gosselin R. (2012). Visible and near-
514 infrared light transmission: A hybrid imaging method for non-destructive meat quality
515 evaluation. *Infrared Physics and Technology*. *55*, 412-420.

516

Table 1. Summary of performances of PLS-1 regression models developed for fat content prediction in ground beef (most accurate models in bold) using dielectric spectral data (ε', ε''and ε''/ε').

		Cal.set (n=35)				Val.set (n=34)		
Frequency	Data type	no. PLS loadings	Slope	R^2CV	RMSECV	R^2P	RMSEP	Bias
ε' (dielectric constant)								
0.6-20GHz	raw	3	0.79	0.73	4.49	0.83	3.11	-0.71
	nor.u.v.	2	0.8	0.77	4.12	0.77	3.64	-0.74
0.6-9.7GHz	raw	2	0.77	0.76	4.2	0.85	2.94	-0.56
	nor.u.v.	3	0.83	0.75	4.32	0.68	4.28	-0.86
12.9-20GHz	raw	2	0.79	0.77	4.1	0.79	3.52	-0.47
	nor.u.v.	**6**	**0.88**	**0.81**	**3.71**	**0.87**	**2.71**	**-0.58**
ε'' (loss factor)								
0.6-20GHz	**raw**	**2**	**0.83**	**0.82**	**3.61**	**0.8**	**3.37**	**-0.37**
	nor.u.v.	5	0.86	0.8	3.84	0.61	4.78	-1.71
0.6-9.7GHz	raw	3	0.83	0.82	3.64	0.77	3.68	-0.34
	nor.u.v.	6	0.84	0.78	4.03	0.59	4.99	-0.07
12.9-20GHz	raw	2	0.82	0.81	3.73	0.8	3.43	-0.43
	nor.u.v.	3	0.72	0.62	5.31	0.49	5.41	-0.99
ε''/ε' (loss tangent)								
0.6-20GHz	raw	4	0.84	0.82	3.64	0.72	4.05	-0.75
	nor.u.v.	2	0.71	0.71	4.63	0.5	5.4	-1.12
0.6-9.7GHz	**raw**	**4**	**0.81**	**0.8**	**3.81**	**0.77**	**3.66**	**-1.27**
	nor.u.v.	3	0.72	0.67	4.93	0.44	5.67	-1.59
12.9-20GHz	raw	2	0.76	0.75	4.28	0.61	4.75	-0.95
	nor.u.v.	3	0.73	0.69	4.79	0.45	5.65	-1.36

PLS-1, partial least squares regression models developed using one Y variable; R^2CV, correlation coefficient of determination in cross-validation; RMSECV, root mean square error of cross-validation; R^2P, coefficient determination of prediction; RMSEP, root mean square error prediction; nor.u.v., normalisation on unit vectors; Cal.set, calibration set; Val.set, validation set.

Table 2. Summary of performances of PLS-1 models developed for fat content prediction in ground beef (most accurate model in bold) using NIR spectra.

NIR (with Martens uncertainty test)		Cal.set n=35				Val.set n=34		
Wavelength range	Data type	no. PLS loadings	Slope	R^2CV	RMSECV	R^2P	RMSEP	Bias
780-2500nm	raw	3	0.97	0.98	1.21	0.98	1.12	-0.22
	nor.u.v.	4	0.99	0.99	1.04	0.98	1.17	0.02
	SNV	3	1	0.98	1.29	0.99	0.9	-0.15
	MSC	3	1	0.98	1.23	0.98	0.98	0.01
1400-2500nm	**raw**	**4**	**1**	**0.99**	**0.84**	**0.99**	**0.71**	**0.04**
	nor.u.v.	3	0.99	0.98	1.23	0.99	0.67	0.01
	SNV	3	0.99	0.98	1.2	0.99	0.78	-0.01
	MSC	3	0.99	0.98	1.2	0.99	0.79	-0.03
900-1800nm	raw	4	0.99	0.98	1.11	0.97	1.24	0.15
	nor.u.v.	4	0.99	0.98	1.09	0.98	1.01	0.15
	SNV	3	0.98	0.97	1.37	0.97	1.22	-0.14
	MSC	2	0.97	0.97	1.45	0.96	1.46	-0.19

PLS-1, partial least squares regression models developed using one Y variable; NIR, near-infrared; R^2CV, correlation coefficient of determination in cross-validation; RMSECV, root mean square error of cross-validation; R^2P, coefficient determination of prediction; RMSEP, root mean square error prediction; MSC, multiplicative scatter correction; SNV, standard normal variate; nor.u.v., normalisation on unit vectors; Cal.set, calibration set; Val.set, validation set.

Table 3. A comparison of PLS-1 and PLS-2 models over 0.6-20 GHz dielectric frequency range for both fat and water content predictions.

	Cal. set n=35					Val. Set n=34		
PLS-1	**PLS loadings**	**Slope**	**R^2CV**	**RMSECV**	**Bias**	**R^2P**	**RMSEP**	**Bias**
Fat	3	0.79	0.73	4.49	-0.1	0.83	3.11	-0.71
Water	3	0.8	0.74	2.79	0.12	0.78	2.24	0.2
PLS-2								
Fat	3	0.79	0.73	4.47	-0.09	0.83	3.17	-0.5
Water	3	0.8	0.74	2.82	0.12	0.78	2.23	0.21

PLS-1, partial least squares regression models developed using one Y variable; PLS-2, partial least squares regression models developed using two Y variables; NIR, near-infrared; R^2CV, correlation coefficient of determination in cross-validation; RMSECV, root mean square error of cross-validation; R^2P, coefficient determination of prediction; RMSEP, root mean square error prediction; Cal.set, calibration set; Val.set, validation set.

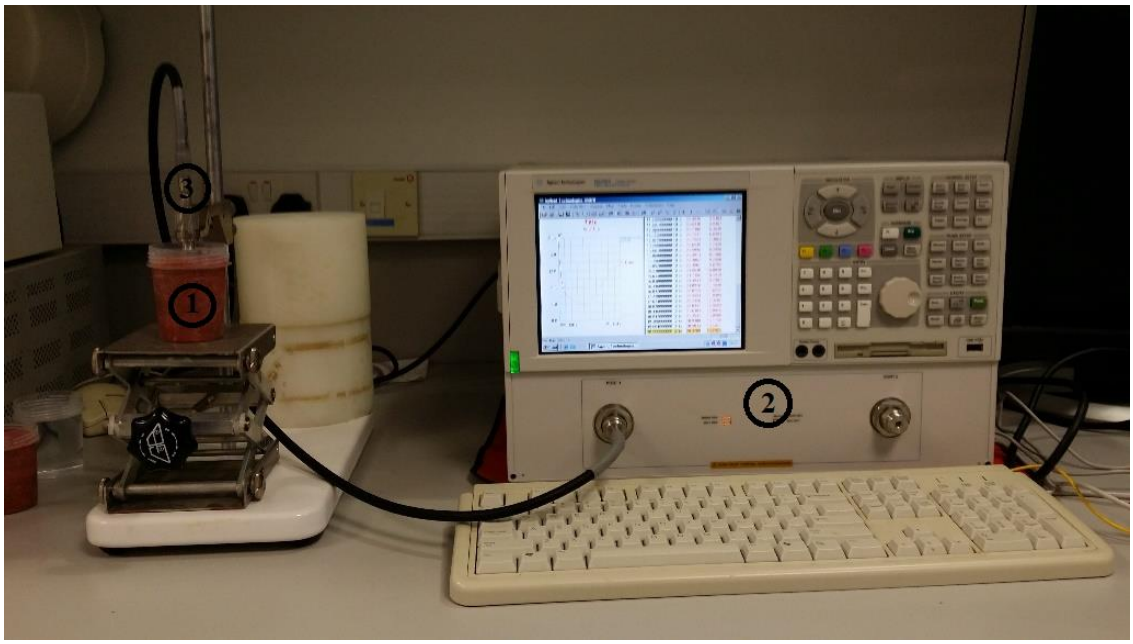


Fig. 1. Measurement of dielectric properties of ground beef samples-① (Agilent 85070E High Temperature coaxial probe kit-③ connected to an Agilent impedance analyser E4991A-② (Agilent technologies, Santa Clara, US)

610

611

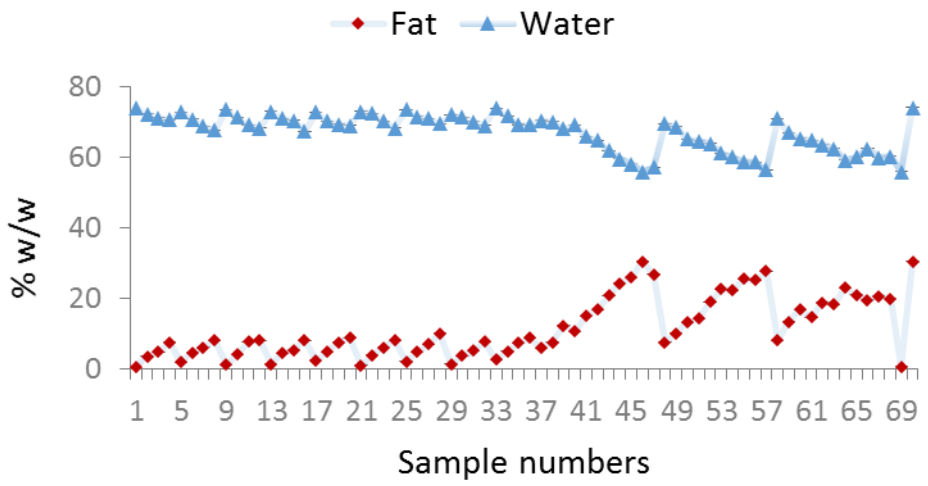
612

613

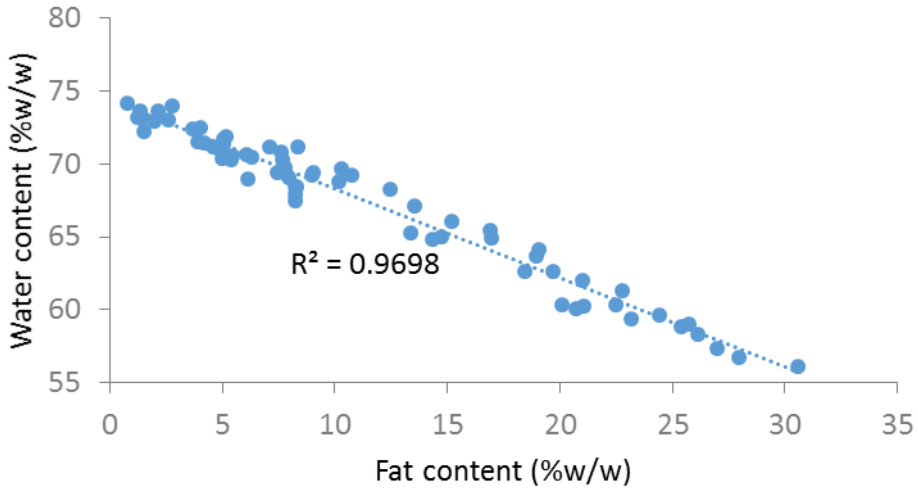
614

615

a.



b.

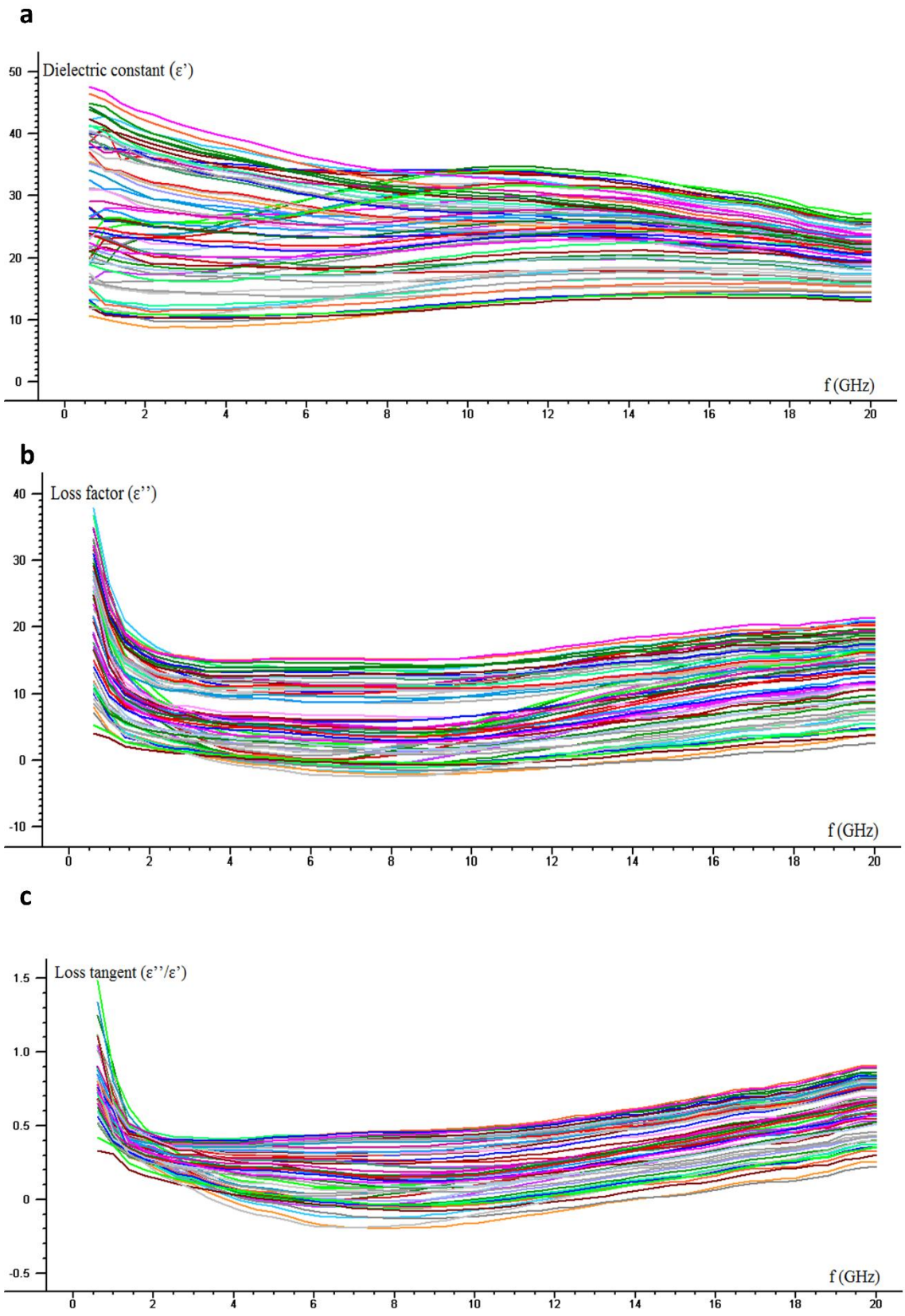


616

617 Fig. 2**a.** fat and water contents (%w/w) of ground beef samples (n=69); **b.** simple linear

618 regression of fat content (%w/w) against water content (%w/w)

619

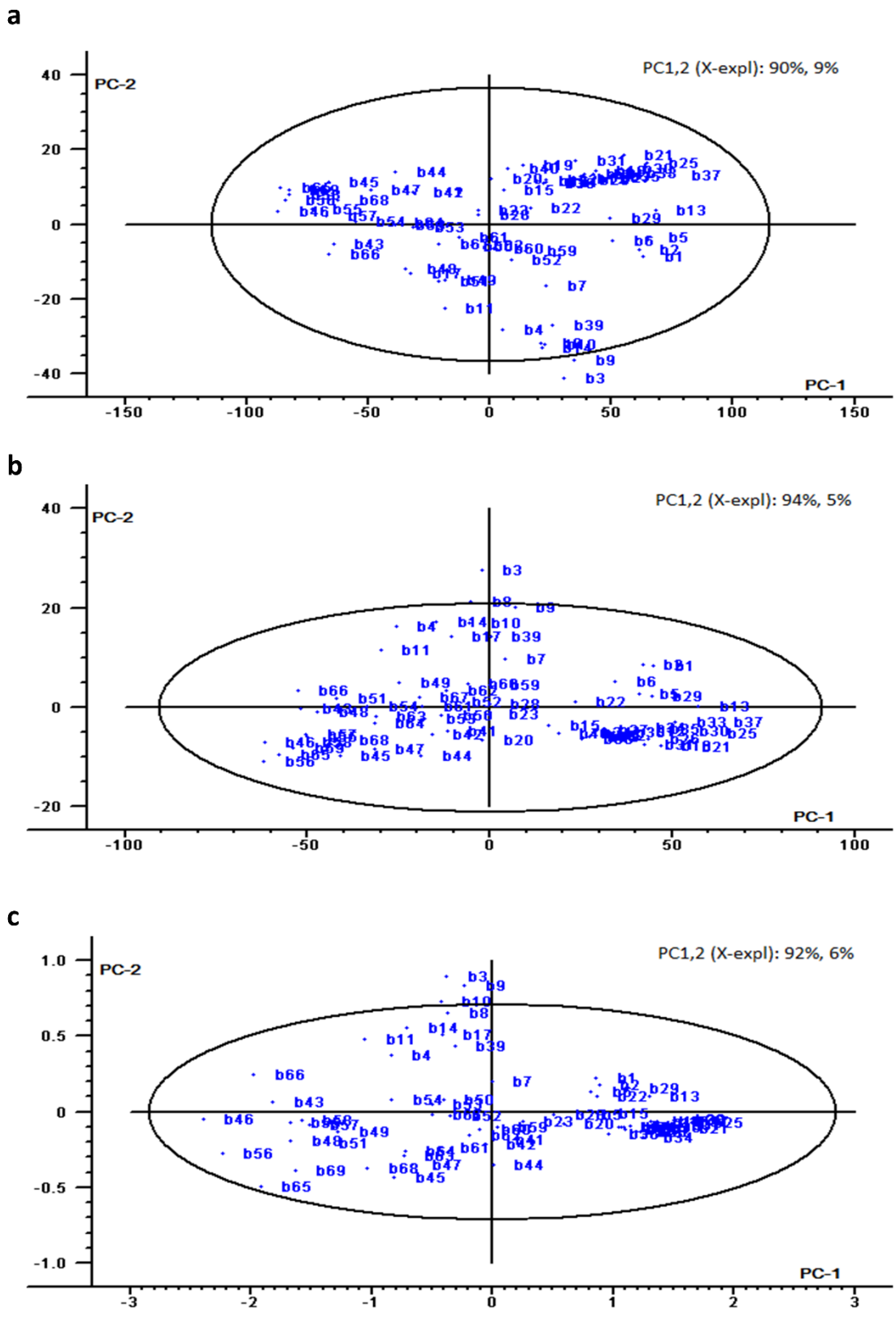


620

621 Fig. 3. Dielectric spectra of **a.** dielectric constant (ε'), **b.** loss factor (ε'') and **c.** loss tangent

622 (ε''/ε') over 0.6-20 GHz frequency range.

623

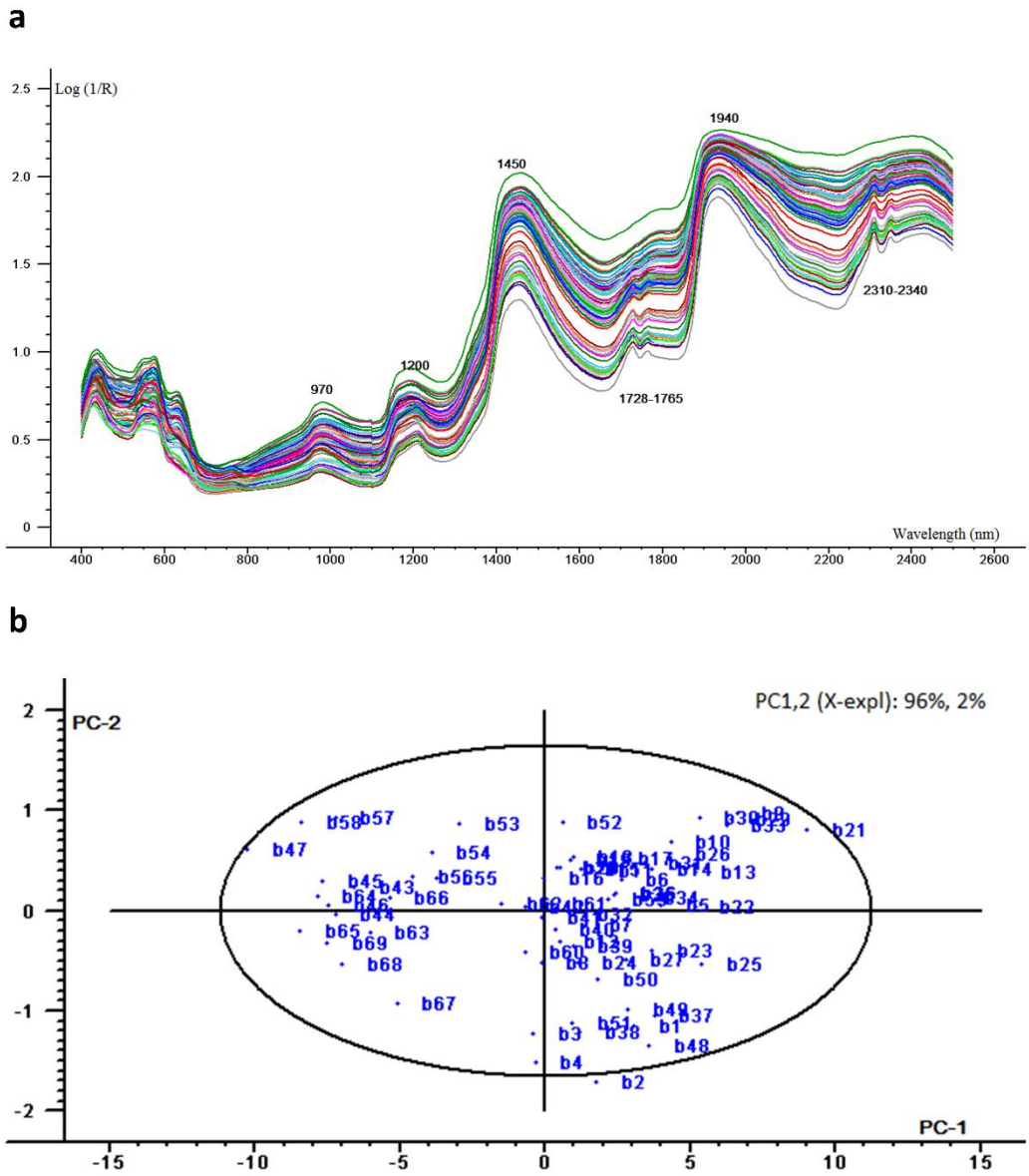


624

625 Fig. 4. PCA score plots of all ground beef samples (n=69) based on spectral data of **a.**

626 dielectric constant (ε'), **b.** loss factor (ε'') and **c.** loss tangent (ε''/ε').

627



628

629 Fig. 5**a.** NIR reflectance (log (1/R)) spectra (400-2498nm) of ground beef samples (n=69); **b.**

630 PCA score plots of all ground beef samples (n=69) based on NIR spectal data (400-2498nm).

631

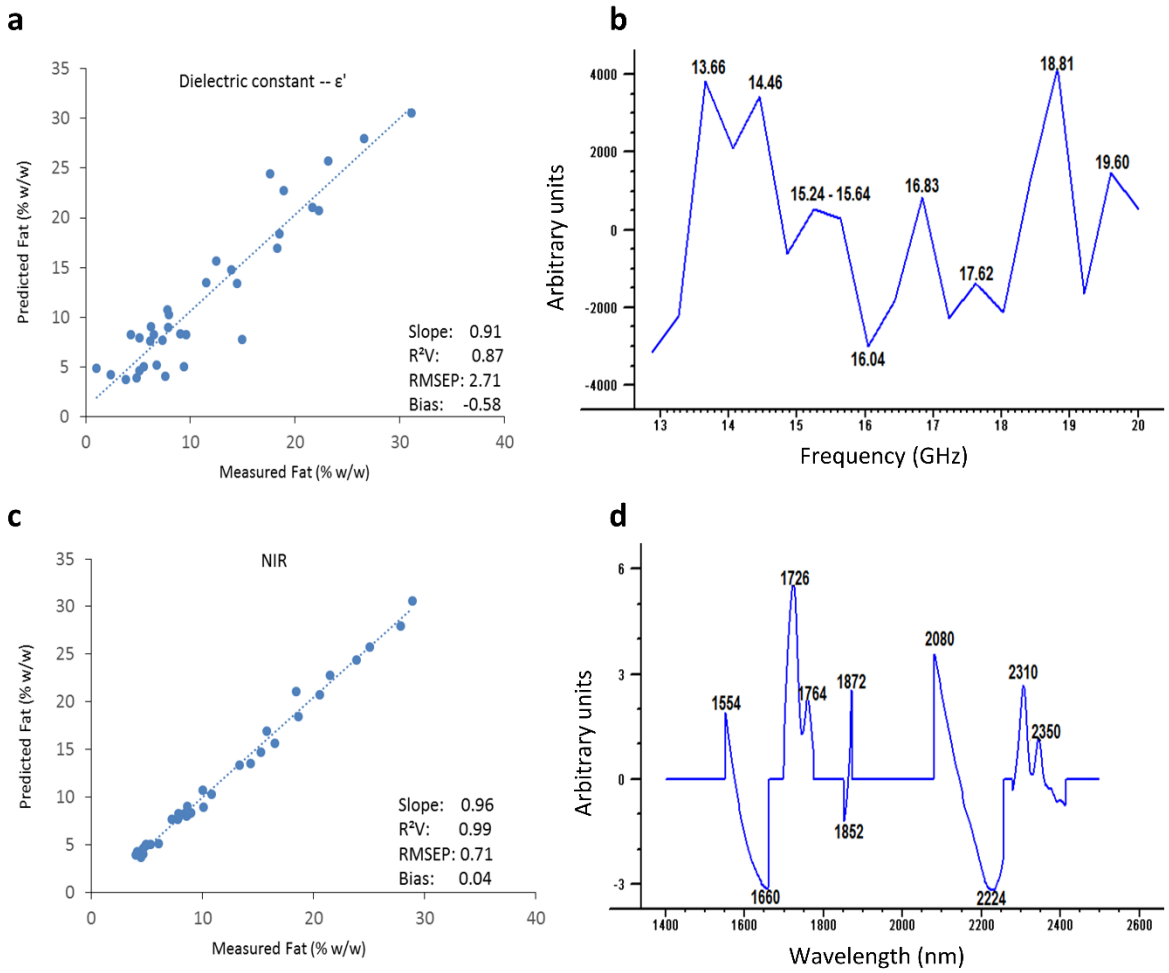
632

633

634

635

636



637

638 Fig. 6. Linear regression plots of fat content (% w/w) reference values versus predicted
639 values of the most accurate model of **a.** dielectric constant--ε' & **c**. NIR. Regression
640 coefficient plots of **b.** dielectric frequency range (12.9-20 GHz) vs. intensity (arbitrary units)
641 for the most accurate model developed using dielectric spectral data (ε'); **d.**wavelength
642 range (1400-2500 nm) vs. intensity (arbitrary units) for the most accurate model developed
643 using NIR spectral data after application of the Martens uncertainty test.

644

645

646

647

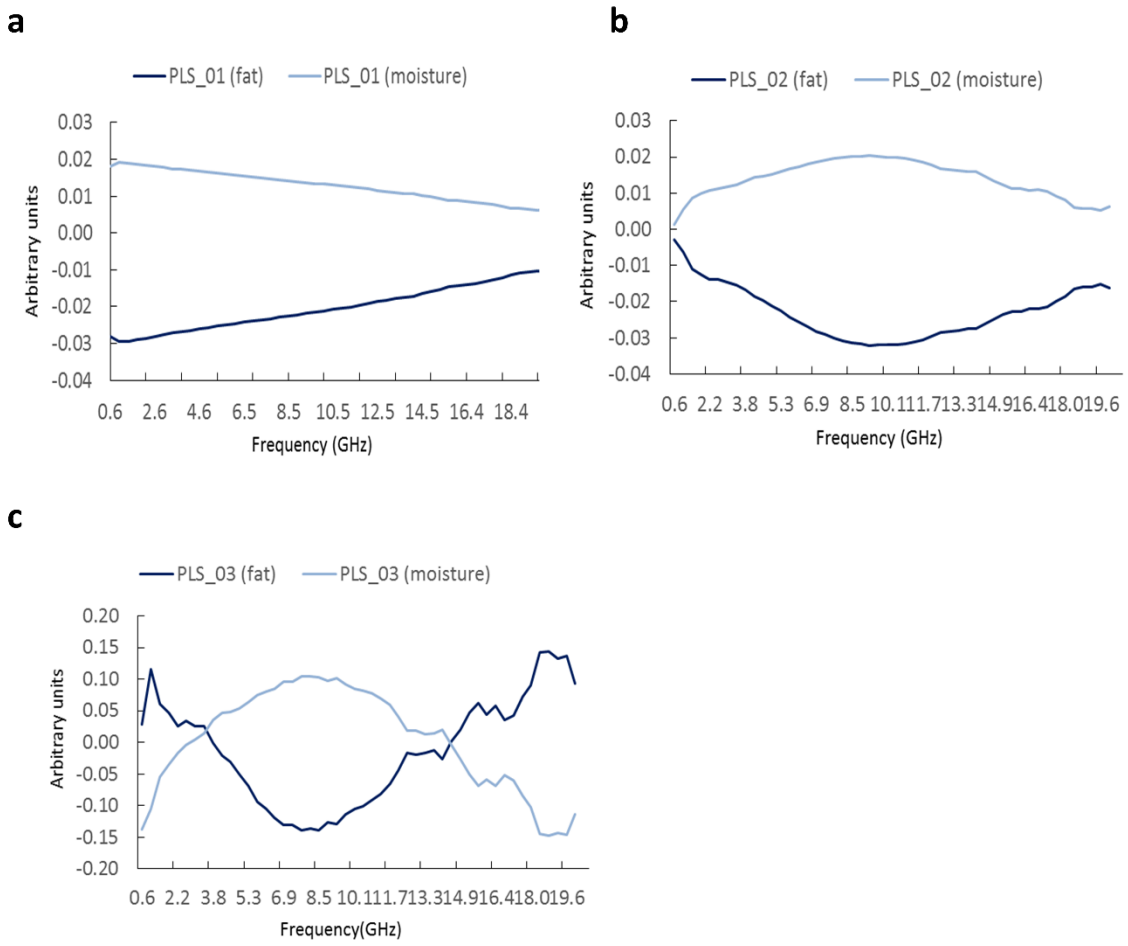


Fig. 7. Regression coefficient plots of the first three PLS vectors [**a.** PLS_01, **b.** PLS_02, **c.** PLS_03] in the PLS-2 modelling for both fat and water content predictions.