Improved by potassium chloride (KCl) dilution predictive ability of attenuated total reflectance Fourier transform infrared (ATR-FTIR) spectroscopy to determine nutrient contents of sunflower meal

S. Yasar[1], M.H. Alma[2], R. Tosun[1,4] and T. Salan[3]

Igdir University
[1] Faculty of Agriculture, Department of Animal Science, 76000, Iğdır, Turkey
[2] Rectorate, 76000, Iğdır, Turkey
[3] Kahramanmaras Sutcu Imam University, Faculty of Forestry, Department of Forest Industrial Engineering, 46100, Kahramanmaras, Turkey

[4] Corresponding author:
e-mail: ramazantosun68@gmail.com

ABSTRACT. Predictive ability of attenuated total reflectance (ATR) Fourier transform (FT) infrared (IR) spectroscopy to determine nutrient contents of sunflower meal (SFM) high in fibre and ash was investigated by testing the effect of potassium chloride (KCl) spiking. The partial least square regression (PLSR) models were generated using original IR spectrum, and its first and second derivate data followed by the normalisation, smoothing and multiplicative scatter correction in order to predict nutrient contents of spiked and non-spiked SFM samples. The results showed that the best model for the prediction was the one derived from the second derivate IR data with high degree of precision and accuracy (R^2 = 0.99, r^2 = 0.95, residual mean square error of cross-validation (RMSECV) = 0.66 for dry matter and 1.30 for crude ash). Furthermore, the precision and accuracy of all three models were improved by the KCl levels spiked at 27.5 and 57.0%, while no effects of further spiking with KCl was observed. In conclusion, the model generated from the second derivate IR data was highly recommended to predict the SFM nutrient contents, and the KCl dilution up to 57% improved the prediction accuracy and precision irrespective to the models used.

Introduction

Sunflower (*Helianthus annuus*) meal (SFM) is by-product of oil extraction from the seeds of sunflower, rich in protein and used for feeding farm animals (Senkoylu and Dale, 1999). Nutritional quality of SFM may vary depending on the proportion of hulls and oils (Ravindran and Blair, 1992; Senkoylu and Dale, 1999). It is comparatively cheap source meeting the protein requirement of farm animals. However, feed industry occasionally needs rapid,

non-destructive, inexpensive and reliable methods determining nutritional quality of SFM products to replace the expensive wet-chemistry methods of analysis.

Attenuated total reflectance (ATR) Fourier transform (FT) infrared (IR) spectral measurement is a fast, non-destructive and inexpensive method, widely used to determine chemical structural changes as well as nutrient contents over a range of food- and feeding-stuffs (Wilson and Belton, 1988; Sivam et al., 2013). Near infrared (NIR) spectroscopy has

been also used to predict nutritional profiles in feeding-stuffs including SFM (Lindberg and Kaila, 1980; Hódsági et al., 2012). Till now, prediction of nutrient contents of SFM by ATR-FTIR spectroscopy has not been reported yet. Quantification of nutritional composition as well as molecular structures of various food and feed materials by multivariate analysis combined with FTIR spectroscopy was presented in several studies, and a partial least square regression (PLSR) has been effectively used (Wilson and Belton, 1988; Ścibisz et al., 2011; Samadi et al., 2013; Li et al., 2015). On the other hand, some variability in spectra measurements caused by instrumental settings, particle size and environmental conditions may lower predictive abilities of a PLSR model. For instance, the ability of model to predict the contents of lignin and holo-cellulose of wood samples high in ash content was found weak due to many overlapped bands at the region of inorganic matter (Pizzo et al., 2015). SFM has high levels of ash (up to 70 g/kg) and fibrous matter (up to 300 g/kg). To separate out these overlapped regions, a non-absorbent and transparent crystal matrix such as potassium chloride (KCl), can be used to spike such samples prior to IR spectra measurements (Theodoridou and Yu, 2013), and this was also used for the samples of liquid or gas states (Jensen et al., 1998; Xu et al., 2018). KCl spiking has another advantage in FTIR studies: the nutrient contents of the SFM samples can be diluted to produce a range of concentrations in order to improve the calibration models by the increased method's precision and accuracy. In general, original IR spectra yield has too broad band at specific regions (water to protein / cellulose to ash). Fine details of overlapped IR regions can be discovered by first (D') or second (D'') derivatives of original IR spectra subjected to data processing such as normalisation by the highest peak (NORM) and smoothing (SMOOT) (Steiner et al., 1972; Wolkers et al., 1998; Yang and Yen, 2002). The effects of such treatments of IR spectra are tested in our study. Furthermore, a test of multiplicative scatter correction (MSC) has been successfully applied to the mean-centred IR spectra in order to prevent any left-over multiplicative and additive effects associated with instrumental, sample and environmental related conditions during the measurements (Isaksson and Næs, 1988; Windig et al., 2008; Li et al., 2015).

Therefore, the aim of this study was to reduce external and internal variability in IR spectra measurement by corrective actions of KCl spiking. For this purpose, PLSR models employing original, first and second derivate IR data were evaluated on the basis of: the number of components explaining the total variance (COMP), a residual sum of squares of cross-validation (RESSCV), a residual mean square error of cross-validation (RMSECV), a standard error of cross validation (SECV), a ratio of performance to the prediction (RPD), regression coefficients of both calibration (R^2) and cross validation (r^2) as well as the precision and accuracy of the method.

Material and methods

In the study, three SFM samples were obtained from a feed mill in Tekirdag (Turkey). The laboratory of the provider confirmed three SFM samples differed in crude protein (CP) contents, and identified as: A – SFM with 36% CP, B – SFM with 34% CP and C – SFM with 28% CP. The contents of CP, crude ash (CA), dry matter (DM) and crude fibre (CF) were chemically re-analysed and cross-validated by our own laboratory using the official method of analysis (AOAC, 1990). SFM samples used in the study contained high levels of CF (up to 25%) and CA (up to 6%). In order to reduce the effects of high levels of CF and CA interfering with IR spectra, the nutrient contents of sample C was diluted by KCl spiking at different levels. This intervention also provided SFM with various concentrations of nutrients, useful in the calibration of the PLSR models. A non-absorbent and transparent matrix, KCl, was used to spike the sample C containing 28% of CP, 6% CA and 25% CF. The sample C was replaced by the KCl (w/w) at 0 (C sample, itself), 27.5 (C1), 57 (C2), 85 (C3) and 100% (C4, pure KCl), whereas the A, B and C samples remained non-spiked.

In total, 7 SFM samples were homogenously ground by a drum mill (Drum-Mill TM 300, Retsch GmbH, Haan, Germany) to reach a final particle size of <20 μm in order to allow sufficient energy pass through the sample during IR measurements. The samples of SFM obtained from the feed mill were carefully sampled from the bunkers into a 5 kg homogenous samples. Homogenous samples were divided into 20 replicate samples at our laboratory and dried. The samples were independently analysed for the contents of DM, CA, CP and CF according to AOAC methods (AOAC, 1990).

Prior to IR measurements, the samples were completely dried out (Yang et al., 2005), and 20 IR measurements were taken from each of 7 samples. All spectra were collected on Agilent Carry 630 FTIR spectrometer with single bounce diamond ATR sampling accessory to ensure highest quality

of spectra (Agilent Technologies Inc., Danbury, CT, USA) at 4 cm⁻¹ resolution in the spectral range of 4000 to 700 cm⁻¹. Each IR measurements included 64 scans which were co-added and corrected for a background spectrum simultaneously. For each SFM, all original 20 IR spectra chosen from the band of 3575 to 880 cm⁻¹ (non-meaningful bands from 2250 to 1900 cm⁻¹ and of >3575 and of <880 cm⁻¹ were cut out) were subjected to first (D') and second derivate (D''). All spectra were consecutively subjected to NORM process by re-scaling according to the highest peak of the region, SMOOT with 9-point used without distorting the peak location and height, and MSC treatments. The data set of 140 spectra, 20 from each of 7 feed samples, was randomly introduced to three PLSR models based on original (model 1), first – D' (model 2) and second – D'' (model 3) derivate spectra in two parts using Minitab statistical programme (Minitab® 17.1.0, Minitab Inc., Coventry, UK). A calibration model was based on the first part (n = 95) after excluding 3 outliers, whilst the cross-validation was on the second part

predictive indicator (Williams and Sobering, 1993; Torres Mariani et al., 2015). In addition, the accuracy calculated as the difference of predicted values and the actual values, the standard deviation (SD) of prediction as a measure of precision within the same sample, R^2 and r^2 were estimated and compared between three models (1, 2 and 3) generated from original, first (D') and second derivate (D'') spectra data.

Results and discussion

Original, D' and D'' IR spectra are shown in Figure 1. It was noted that the number of bands was equal to the derivative order number plus one with the change of spectra from the negative slope to the positive slope. Increasing degree of derivatisation has been reported to result in more distinguishable derivatives (Whitaker and Pigford, 1960; Morrey, 1968). The studies reported better prediction parameters with the models using D'' spectra than the models using original and D' spectra (Isaksson and
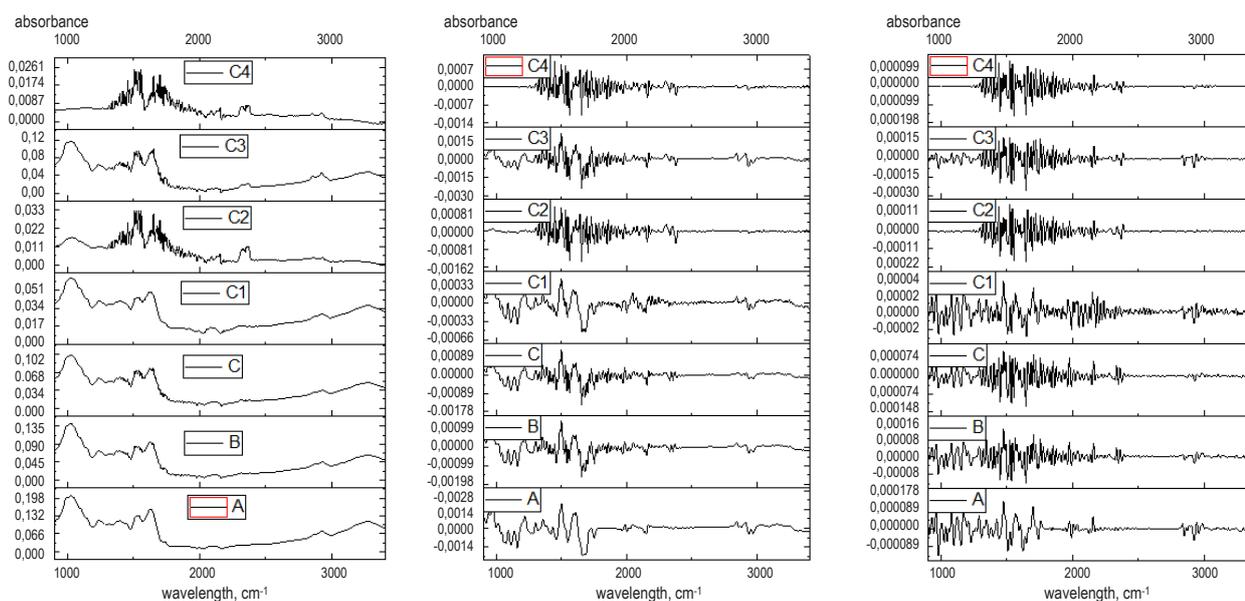

Figure 1. Original (left), D' (middle) and D'' (right) of IR spectra (over 3500 to 800 cm⁻¹ band region) of SFM samples differing in nutrient contents. Samples: A, B and C indicate the SFM with 36, 34 and 28% CP contents, C1, C2 and C3 indicate the spiking levels with KCl at 27.5, 57.0 and 85.0%, respectively, and C4 is the IR spectra of the pure sample of KCl.

(n = 42 as unknown) using a leave-one-out cross validation and predicting as unknown samples. This continues until a minimum RESSCV and RMSECV were found. The model was then applied to the cross validation data set to predict the nutrient concentrations of 7 SFM samples. The SECV was then estimated. The RPD was also calculated; the values higher than 2 were considered as good

Næs, 1988; Windig et al., 2008; Zhang et al., 2015). Similarly, we found that the D'' spectra data was also capable of establishing a strong linear relationship between the chemical composition and the peaks over the IR band region (from 3500 to 800 cm⁻¹). Herein, the performance characteristics of our PLSR models were presented and discussed.

Table 1. The power of partial least square regression (PLSR) models used to predict nutrient contents from the absorbance peaks over a infrared IR spectral range from 3500 to 800 cm^{-1} (97 data sets used in the calibration stage, 42 data sets in the cross validation stage, 3 outliers)

Models	Nutrients	COMP	V, %	RESSCV	RMSECV	R^2	r^2
Original	DM	7	99.66	105.71	1.93	0.91	0.80
(model 1)	CA			7.83	0.59	0.98	0.96
	CP			591.42	1.78	0.95	0.93
	CF			465.23	1.50	0.95	0.89
D'	DM	4	82.88	141.94	1.12	0.92	0.85
(model 2)	CA			18.43	0.71	0.95	0.92
	CP			1096.21	1.77	0.95	0.87
	CF			660.63	1.71	0.92	0.85
D"	DM	9	86.54	33.08	0.66	0.99	0.95
(model 3)	CA			121.20	1.30	0.99	0.94
	CP			339.19	0.99	0.99	0.94
	CF			185.56	0.90	0.99	0.96

PLSR analysis of variance of DM – dry matter, CA – crude ash, CP – crude protein and CF – crude fibre were significant at $P < 0.000$; COMP – the number of components; V – the percentage of total variance explained by the components; RESSCV – residual sum of squares of cross-validation; RMSECV – residual mean square error of cross-validation

In terms of performance parameters, the model 1 (original spectra) enabled to explain 99.66% of variance by 7 components, while the model 2 (first derivate) explained 82.88% of variance by 4 components, and the model 3 (second derivate) explained 86.54% of variance by 9 components (Table 1). All models explained sufficient amount of variance. But, the model 3 produced the lowest values of RESSCV (33.08 for DM and 121.20 for CA) and RMSECV (0.66 for DM and 1.30 for CA). The model 1 produced a middle range of value for RESSCV (7.73 for CA and 591.42 for CP) and RMSECV (0.59 for CA and 1.93 for DM) values, while the model 2 has the highest RESSCV (18.43 for CA and 1096.21 for CP) and RMSECV (0.71 for CA and 1.77 for CP) values. The highest R^2 value (0.99) was obtained from the model 3 in comparison to the values of model 1 (0.92–0.98) and model 2 (0.91–0.95). The model with higher r^2 values can be a good model. Apparently, the highest r^2 values were obtained from the model 3 (0.94 to 0.96) as compared to the r^2 values of model 1 (0.85 to 0.96) and model 2 (0.80 to 0.92). Similar prediction parameters were previously reported with the D" of spectra data (Zhang et al., 2015). This is simply due to 'the more distinguishable peaks extracted by second derivatisation process the lower random errors into a calibration model' (Bjørsvik and Martens, 2007).

The chemical analysis of SFM samples is presented in Tables 2 and 3. Spiking C sample with KCl produced a good range of nutrient concentra-tions. The respected models derived from the PLSR analysis which regressed the absorbance peaks over IR spectra ranging from 3500 to 800 cm^{-1} with the concentrations of nutrients in SFM produced the following prediction equations:

$$Y = b_0 + b_1X_1 + b_2X_2 + ... + b_pX_p,$$

where: Y – concentration of nutrient; b_0 – regression coefficient for the intercept; b_i – regression co-efficients (for variables 1 through p) computed from the absorbance peaks over a IR spectra (2350 points) ranging from 3500 to 800 cm^{-1}. For the sake of clarity, herein the equations for the entire grading range under study were summarised in the following linear regression equations derived from regressing the predicted values with the analysed values.

Equations of the model 1 using original, D' and D" spectra data predicting DM of SFM were as follows:
$y = 0.8192x + 16.594$ and $R^2 = 0.7936$ (original),
$y = 0.8274x + 15.903$ and $R^2 = 0.8698$ (D'), and
$y = 1.0616x - 5.6395$ and $R^2 = 0.972$ (D"), respectively.

Equations of the model 1 using original, D' and D" spectra data predicting CA of SFM were as follows:
$y = 0.9249x + 0.4427$ and $R^2 = 0.9125$ (original),
$y = 0.8611x + 0.7602$ and $R^2 = 0.8801$ (D'), and
$y = 0.7692x - 1.1778$ and $R^2 = 0.7476$ (D"), respectively.

Equations of the model 1 using original, D' and D" spectra data predicting CP of SFM were as follows:
$y = 0.9194x + 2.4008$ and $R^2 = 0.8639$ (original),
$y = 0.9083x + 2.6542$ and $R^2 = 0.9334$ (D'), and
$y = 0.9566x - 1.4304$ and $R^2 = 0.9867$ (D"), respectively.

Equations of the model 1 using original, D' and D" spectra data predicting CF of SFM were as follows:
$y = 0.8415x + 3.5966$ and $R^2 = 0.7912$ (original),
$y = 0.8387x + 3.4953$ and $R^2 = 0.8680$ (D'), and
$y = 0.8845x - 2.7486$ and $R^2 = 0.9270$ (D"), respectively.

Tables 2 and 3 showed the results of the predicted means by PLSR models. In the light of SECV, RPD, precision and accuracy values (Table 2), the model 3 has a highest predictive ability for DM contents than the models 1 and 2. The regression coefficients of R^2 and r^2 of 0.99 and 0.95 of the model 3 were higher than the corresponding coefficient values of 0.92 and 0.85 of the model 1, and 0.90 and 0.80 of the model 2 (Table 1). Furthermore, these values were well lined up with other parameters. Namely, the RMSECV and RESSCV values presented in Table 1 were gradually lowered by

the PLSR model 3, compared to the PLSR models of 1 and 3, respectively (Table 2). As shown in Table 2, the highest RPD (3.21 to 3.26) for DM prediction of non-spiked SFM samples were obtained from the PLSR model 3, compared the RPD values from the model 1 (1.56 to 1.67) and model 2 (1.41 to 1.75). Spiking SFM sample with KCl at a rate of 27.5% (sample C1) improved the predictive abilities of all PLSR models for DM contents. The RPD values in all models were improved by nearly 2-fold by 27.5% KCl spiking. The RPD values of other spiked levels were out range of the prediction. In other words, further spiking with KCl >27% did not improve, but worsened the predictive ability for the DM.

Table 2. Predictive abilities of partial least square regression (PLSR) models for dry matter (DM) and crude ash (CA) content in sunflower meal (SFM) from the absorbance peaks over a infrared (IR) spectral range from 3500 to 800 cm⁻¹

Indices	Non-spiked SFM samples[1]			Sample C spiked with KCl[1]			
	A	B	C	C1	C2	C3	C4
DM analysed values							
mean, %	88.84	91.37	90.61	92.94	96.13	99.00	99.23
SD, %	0.49	0.55	0.45	0.89	0.18	0.36	0.07
Predicted values by the model 1 using original spectra data							
mean, %	90.69	89.57	90.87	92.40	96.40	95.19	100.17
SD, %	0.332	0.299	0.302	0.355	0.66	0.87	0.21
SECV	0.292	0.358	0.287	0.392	0.618	1.160	0.521
RPD	1.67	1.53	1.56	2.27	0.29	0.310	0.00
Predicted values by the model 2 using D' spectra data							
mean, %	90.01	89.36	92.1	92.53	96.12	96.44	99.30
SD, %	0.86	0.53	0.51	0.40	0.22	1.52	0.27
SECV	0.28	0.39	0.26	0.33	0.33	0.53	0.40
RPD	1.75	1.41	1.73	2.69	0.54	0.68	0.17
Predicted values by the model 3 using D" spectra data							
mean, %	89.10	90.70	90.96	92.80	95.84	100.76	99.00
SD, %	0.84	0.45	0.33	0.20	0.14	1.50	0.21
SECV	0.15	0.17	0.14	0.14	0.14	0.49	0.16
RPD	3.26	3.23	3.21	6.35	1.00	0.73	0.00
CA analysed values							
mean, %	4.99	6.06	4.86	4.52	3.30	0.90	0.0
SD, %	0.50	0.30	0.62	0.20	0.89	0.30	0.0
Predicted values by the model 1 using original spectra data							
mean, %	5.28	5.80	5.27	4.56	2.90	2.42	−0.35
SD, %	0.17	0.10	0.07	0.14	0.20	0.34	0.54
SECV	0.088	0.10	0.086	0.11	0.18	0.350	0.15
RPD	6.25	3.00	7.20	1.81	4.94	0.46	0.00
Predicted values by the model 2 using D' spectra data							
mean, %	5.73	5.93	4.77	4.67	2.60	2.66	0.17
SD, %	0.48	0.30	0.31	0.23	0.11	0.98	0.14
SECV	0.12	0.15	0.10	0.13	0.13	0.17	0.16
RPD	4.16	2.00	6.20	1.53	6.84	1.76	0.00
Predicted values by the model 3 using D" spectra data							
mean, %	5.14	6.02	4.75	4.55	3.05	3.83	−0.15
SD, %	0.16	0.08	0.23	0.12	0.11	1.00	0.42
SECV	0.086	0.090	0.076	0.075	0.073	0.263	0.087
RPD	5.81	3.33	8.18	2.66	12.2	1.14	0.00

[1] sunflower meal samples: A (SFM-36-CP), B (SFM-34-CP), C (SFM-28-CP), C1 (C spiked with KCl at 27.5%), C2 (C spiked with KCl at 57%), C3 (C spiked with KCl at 85%) and C4 (pure KCl); SD – standard deviation; SECV – standard error of cross validation; RPD – a ratio of performance to the prediction; D' − first derivate; D" − second derivate

Table 3. Predictive abilities of partial least square regression (PLSR) models for the percentage (%) of crude protein (CP) and crude fibre (CF) in sunflower meal (SFM) from the absorbance peaks over a infrared (IR) spectral range from 3500 to 800 cm⁻¹

Indices	Non-spiked SFM samples[1]			C sample spiked with KCl[1]			
	A	B	C	C1	C2	C3	C4
CP analysed values							
mean, %	36.03	34.07	28.02	20.1	12.00	4.03	0.0
SD, %	1.05	1.07	1.04	1.01	0.90	0.02	0.0
Predicted values by the model 1 using original spectra data							
mean, %	31.04	36.32	30.15	22.76	8.86	14.67	−3.57
SD, %	1.37	1.24	1.12	1.66	3.07	3.02	3.80
SECV	0.79	0.97	0.78	1.06	1.68	3.14	1.41
RPD	1.32	1.10	1.33	0.95	0.53	0.006	0.00
Predicted values by the model 2 using D' spectra data							
mean, %	33.78	37.8	24.84	22.27	10.15	11.18	0.50
SD, %	3.23	1.90	1.57	1.41	0.71	6.33	1.19
SECV	0.78	0.97	0.60	0.83	0.83	1.09	1.00
RPD	1.34	1.10	1.77	1.21	1.08	0.018	0.00
Predicted values by the model 3 using D" spectra data							
mean, %	35.76	35.82	26.25	20.52	12.25	7.71	0.12
SD, %	2.31	1.23	0.57	1.00	0.71	1.34	1.22
SECV	0.33	0.37	0.33	0.32	0.31	1.123	0.38
RPD	3.18	2.89	3.15	3.15	3.00	0.01	0.00
CF analysed values							
mean, %	21.22	23.30	25.30	18.48	10.97	3.83	0.0
SD, %	1.01	1.02	1.00	0.90	0.50	0.04	0.0
Predicted values by the model 1 using original spectra data							
mean, %	22.67	24.65	22.60	19.02	10.11	14.81	−1.92
SD, %	0.59	0.63	0.66	0.74	2.00	2.86	2.13
SECV	0.28	0.32	0.27	0.26	0.25	0.92	0.31
RPD	3.67	3.18	3.73	3.46	2.00	0.04	0.00
Predicted values by the model 2 using D' spectra data							
mean, %	24.31	25.05	20.10	19.70	10.68	10.61	0.49
SD, %	2.09	1.34	1.42	1.04	0.52	4.19	0.59
SECV	0.66	0.90	0.60	0.78	0.77	1.03	0.93
RPD	1.53	1.13	1.70	1.28	0.65	0,04	0.00
Predicted values by the model 3 using D" spectra data							
mean, %	23.28	22.86	23.4	20.08	10.98	10.25	−0.42
SD, %	1.11	0.21	0.72	0.21	0.33	1.14	1.63
SECV	0.28	0.32	0.27	0.26	0.25	0.92	0.31
RPD	3.60	3.18	3.70	3.46	2.00	0.04	0.00

[1] sunflower meal samples: A (SFM-36-CP), B (SFM-34-CP), C (SFM-28-CP), C1 (C spiked with KCl at 27.5%), C2 (C spiked with KCl at 57%), C3 (C spiked with KCl at 85%) and C4 (pure KCl); SD – standard deviation; SECV – standard error of cross validation; RPD – a ratio of performance to the prediction; D' − first derivate; D" − second derivate

In contrary, the calculated precision and accuracy values in Figure 2 revealed that the spiking levels of 27.5 and 57% (samples C1 and C2) were found to have a great improvement effect on the prediction abilities of all models, whilst there was no effect of spiking at 85%.

The models, the best model for the CA prediction of non-spiked samples was the model 3, from which the lowest SECV, RESSCV and RMSECV values and the highest R^2, r^2 and RPD values were obtained (Table 2). However, both the model 1 and 2 performed better for the CA than the DM, CP and CF. As overall, the SECV and RPD values were too low in three models for the CA in comparison to the values obtained from the non-spiked samples, except the fact that the RPD value of the spiked sample at 57% of KCl was significantly higher than the remaining spiked and non-spiked samples in all models.
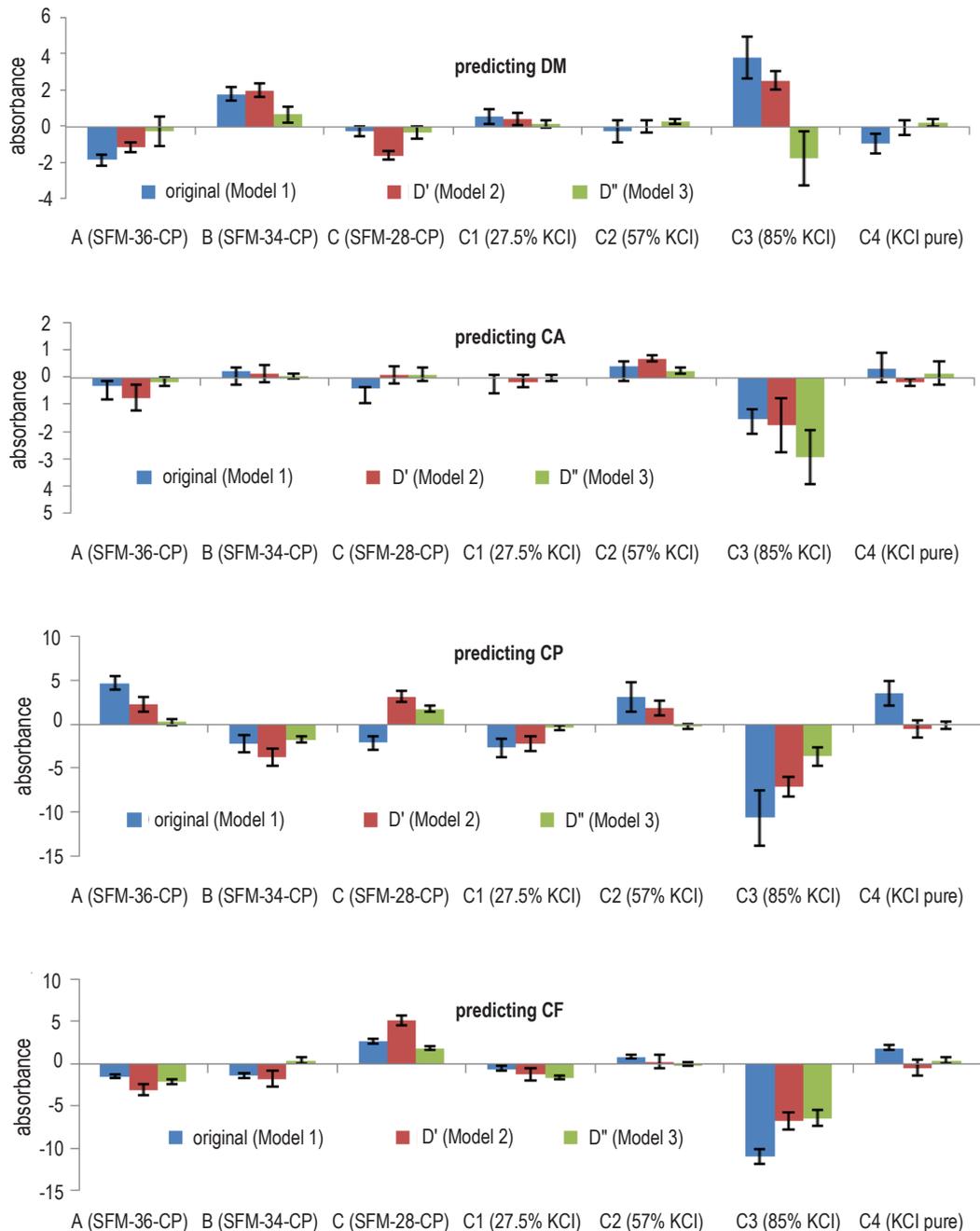

Figure 2. Precision (error of the prediction) and accuracy (refers to the closeness which must be '0' if the predicted value is the same as analysed value) of three partial least square regression (PLSR) models for (A) dry matter (DM), (B) crude ash (CA), (C) crude protein (CP) and (D) crude fibre (CF) (each data point represents the mean with standard error of prediction calculated from n = 20 spectra analysis of each sample)

Samples: A, B and C indicate the sunflower meal (SFM) samples with 36, 34 and 28% crude protein contents; C1, C2 and C3 indicate the spiking levels with KCl at 27.5, 57 and 85%, respectively; and C4 is the infrared (IR) spectra of the pure sample of KCl

On the other hand, the results of precision and accuracy values of three models to predict CA in spiked samples clearly indicated that the spiking at 27.5% has a significant effect to improve the prediction abilities of all models, while the other spiking levels had no observed effect.

For the prediction of CP, the model 3 performed better than the model 1 and model 2 (Tables 1 and 3). The RMSECV value in model 3 was lower than the values in models 1 and 2 (0.99 vs 1.78 and 1.77, respectively). Similarly, the values of R^2 and r^2 of the model 3 were higher than the values of the model 1 and model 2 (0.99–0.95 vs 0.95–0.93 and 0.95–0.87, respectively). The lowest SECV values (0.33–1.123) were determined for the model 3 as compared to the values of model 1 (0.78–3.14) and model 2 (0.60–1.09). The RPD values of the model 1 and model 2 were significantly lower than the value of 2.0, while the RPD values of the model 3 were in the range of 2.89 to 3.18 of non-spiked samples. More surprisingly, there were no effects of spiking on the degree of prediction abilities of all the models according to the RPD values given in Table 3. However, the Figure 2 shows that only the prediction ability of the model 3 was improved by the use of spiked samples at 27.5 and 57.0% of KCl, while no effects of spiking on the prediction ability were seen in the model 1 and model 2.

For the prediction of CF, the model 3 was found significantly better than the model 1 and model 2 (Tables 1 and 3). The RMSECV value in model 3 (0.99) was lower than the values in model 1 (1.50) and in model 2 (1.71). Higher values of R^2 (0.99) and r^2 (0.94) were found in the model 3 in comparison to the values of the model 1 (0.95 and 0.89) and model 2 (0.92 and 0.85). The SECV values (0.27–0.92) of the model 3 were most likely identical to the SECV values of model 1 (0.27–0.92), while the SECV values of the model 2 were higher (0.60–1.03). The RPD values of the model 1 and model 3 were higher than the value of 3.0, while the RPD values of the model 2 was <2.0. The RPD values of spiked samples in all three models were not changed, but even lowered by increasing spiked levels. In contrary, the precision and accuracy of all three models were significantly improved by the use of spiked samples at 27.5 and 57.0% of KCl.

These results indicated that nutrient concentrations of SFM samples were consistently well predicted by the model 3 using second derivate of IR spectra data. The prediction parameters of r^2 (0.96–0.98) and RSECV (1.1–0.18) of CA, CP and CF of the SFM samples by the PLSR model using the second deri-

vate of NIR spectra (Perten, 2019) were similar to our results. In the study of Mahesar et al. (2011) the first derivate of IR spectra data produced an excellent PLSR model to predict the fatty acid contents of poultry lipids. In that study, the first derivate of IR spectra band, specifically assigned to lipid region, was used to establish a powerful model, while the models used in our study employed the entire IR region. Data processing of original IR spectra and their first and second derivates have now routinely been used in IR spectroscopy studies. Similar to our results, considerably low prediction errors and high correlation coefficients with PLSR model 3 were observed by Isaksson and Næs (1988), Windig et al. (2008), Ribeiro et al. (2013), Zhang et al. (2015) and Hell et al. (2016). Comparing the results of PLSR analysis of NIR and FTIR spectra to predict the CA, DM, CP and CF contents of wheat bran high in ash and fibre, FTIR was found superior than NIR to predict the level of CP in the feed materials; although the NIR predicted better CA and crude fibre (CF) contents (Hell et al., 2016).

Many studies reported that the use of FTIR spectra data for nutrient quantification studies would provide more information than the use of NIR spectra data since the FTIR spectra provide also a chemical and molecular structural data (Wilson and Belton, 1988; Shiroma and Rodriguez-Saona, 2009; Ścibisz et al., 2011; Xu et al., 2018). However, both methods of prediction were poor to predict soluble dietary fibre content. It was also pointed out the difficulty of establishment of a good prediction model in such food materials (Pizzo et al., 2015). However, spiking SFM (high in ash and fibre contents) in KCl gradually overcome this problem in our study. Furthermore, FTIR spectroscopy detects fundamental molecular vibrations as opposed to the overlapping and usually weaker combination bands in NIR. The use of KCl to spike the SFM sample high in ash and fibre was indeed very helpful to enhance the detection of molecular vibrations, and thereby to improve the predictive abilities of PLSR models, and this was also confirmed by previous results (Theodoridou and Yu, 2013; Pizzo et al., 2015).

The model 3 using the second derivate of IR spectra data was more robust, having an RPD value of >2.5 as well as better precision and accuracy, and was the best predicting model for all nutrient contents of SFM samples used in our study. Similar results with the model 3 was also reported previously (Zhang et al., 2015), but the values of RPD were closely related to the values of bias (accuracy), as opposed to the values obtained from the spiked SFM samples in our study. Accuracy (closeness of the predicted value

to the analysed value) and precision (degree of standard error in prediction) of the predicted nutrient contents of non-spiked SFM samples were consistently better by the model 3 than the model 1 and model 2 of PLSR in this study. There was a remarkable improvement in the precision and accuracy of all three PLS models by spiking the SFM sample at the rate of 27.5% for the DM and CF determinations, whereas only the level of 27.5% spiking has an improving effect on the degree of precision and accuracy of all the PLSR prediction models for the CA determination. In contrary, both the level of 27.5 and 57.0% spiking have a great improvement effect on the precision and accuracy of only the model 3 for the CP determination in this study. The remaining levels of spiking did not affect the precision and accuracy of the models used in this study (Figure 2). Our results demonstrated that the prediction ability of methods for nutrient determination can be better estimated by the method precision and accuracy values than the values of RPD, which are the ratio of the standard error of performance and the standard deviation of analysed data (Williams and Sobering, 1993; Torres Mariani et al., 2015). However these values cannot provide any information on the accurate and precise. Furthermore, there are no close relationship between the RPD values and the regression coefficients (R^2 and r^2) of the models. Therefore, the 'accuracy' and 'precision in measurement' must be reported in IR measurements.

Conclusions

The nutritional qualities of sunflower meal (SFM) samples are precisely predicted by attenuated total reflectance Fourier transform infrared (ATR-FTIR) spectroscopy using second derivate IR data in comparison to original and first derivate of IR data. What is more, the prediction accuracy of nutrient contents was significantly improved by the use of KCl as diluent.

References

AOAC, 1990. Official Methods of Analysis of the Association of Official Analytical Chemists. 15th Edition. Arlington, VA (USA)

Bjørsvik H.-R., Martens H., 2007. Data analysis: calibration of NIR instruments by PLS regression. In: D.A. Burns, E.W. Ciurczak (Editors): Handbook of Near-Infrared Analysis. 3rd Edition. CRC Press, Taylor & Francis Group, Boca Raton, FL (USA), pp.189–206

Hell J., Prückler M., Danner L., Henniges U., Apprich S., Rosenau T., Kneifel W., Böhmdorfer S., 2016. A comparison between near-infrared (NIR) and mid-infrared (ATR-FTIR) spectroscopy for the multivariate determination of compositional properties in wheat bran samples. Food Control 60, 365–369, https://doi.org/10.1016/j.foodcont.2015.08.003

Hódsági M., Gergely S., Gelencsér T., Salgó A., 2012. Investigations of native and resistant starches and their mixtures using near-infrared spectroscopy. Food Bioprocess Technol. 5, 401–407, https://doi.org/10.1007/s11947-010-0491-5

Isaksson T., Næs T., 1988. The Effect of multiplicative scatter correction (MSC) and linearity improvementin NIR spectroscopy. Appl. Spectrosc. 42, pp. 1273-1284, https://doi.org/10.1366/0003702884429869

Jensen A., Dam-Johansen K., Wójtowicz M.A., Serio M.A., 1998. TG-FTIR study of the influence of potassium chloride on wheat straw pyrolysis. Energy Fuels 12, 929–938, https://doi.org/10.1021/ef980008i

Li J., Danao M.-G.C., Chen S.-F., Li S., Singh V., Brown P.J., 2015. Prediction of starch content and ethanol yields of sorghum grain using near infrared spectroscopy. J. Near Infrared Spectrosc. 23, 85–92, https://doi.org/10.1255/jnirs.1146

Lindberg J.J., Kaila T., 1980. NIR and spin probe studies on amylose and amylopectin. Acta Chem. Scand. B 34, 757–761, https://doi.org/10.3891/acta.chem.scand.34b-0757

Mahesar S.A., Sherazi S.T.H., Kandhro A.A., Bhanger M.I., Khaskheli A.R., Talpur M.Y., 2011. Evaluation of important fatty acid ratios in poultry feed lipids by ATR FTIR spectroscopy. Vib. Spectrosc. 57, 177–181, https://doi.org/10.1016/j.vibspec.2011.06.009

Morrey J.R., 1968. On determining spectral peak positions from composite spectra with a digital computer. Anal. Chem. 40, 905–914, https://doi.org/10.1021/ac60262a006

Perten, 2019. Perten Instruments Application Note. Analysis of Sunflower Meal Using the DA 7250 Analyzer. Perten, Hägersten (Sweden), https://www.perten.com/Global/Application%20notes/DA%2072xx/Perten%20DA%207250%20Sunflower-seed%20meal.pdf

Pizzo B., Pecoraro E., Alves A., Macchioni N., Rodriguez J.C., 2015. Quantitative evaluation by attenuated total reflectance infrared (ATR-FTIR) spectroscopy of the chemical composition of decayed wood preserved in waterlogged conditions. Talanta 131, 14–20, https://doi.org/10.1016/j.talanta.2014.07.062

Ravindran V., Blair R., 1992. Feed resources for poultry production in Asia and the Pacific. II. Plant protein sources. Worlds Poult. Sci. J. 48, 205–231, https://doi.org/10.1079/WPS19920017

Ribeiro L.F., Peralta-Zamora P.G., Maia B.H.L.N.S., Ramos L.P., Pereira-Netto A.B., 2013. Prediction of linolenic and linoleic fatty acids content in flaxseeds and flaxseeds flours through the use of infrared reflectance spectroscopy and multivariate calibration. Food Res. Int. 51, 848–854, https://doi.org/10.1016/j.foodres.2013.01.061

Samadi, Theodoridou K., Yu P., 2013. Detect the sensitivity and response of protein molecular structure of whole canola seed (yellow and brown) to different heat processing methods and relation to protein utilization and availability using ATR-FT/IR molecular spectroscopy with chemometrics. Spectrochim. Acta A Mol. Biomol. Spectrosc. 105, 304–313, https://doi.org/10.1016/j.saa.2012.11.096

Senkoylu N., Dale N., 1999. Sunflower meal in poultry diets: a review. Worlds Poult. Sci. J. 55, 153–174, https://doi.org/10.1079/WPS19990011

Shiroma C., Rodriguez-Saona L., 2009. Application of NIR and MIR spectroscopy in quality control of potato chips. J. Food Comp. Anal. 22, 596–605, https://doi.org/10.1016/j.jfca.2008.09.003

Sivam A.S., Sun-Waterhouse D., Perera C.O., Waterhouse G.I.N., 2013. Application of FT-IR and Raman spectroscopy for the study of biopolymers in breads fortified with fibre and polyphenols. Food. Res. Int. 50, 574–585, https://doi.org/10.1016/j.foodres.2011.03.039

Steiner J.,Termonia Y., Deltour J., 1972. Smoothing and differentiation of data by simplified least square procedure. Anal. Chem. 44, 1906–1909, https://doi.org/10.1021/ac60319a045

Ścibisz I., Reich M., Bureau S., Gouble B., Causse M., Bertrand D., Renard C.M.G.C., 2011. Mid-infrared spectroscopy as a tool for rapid determination of internal quality parameters in tomato. Food Chem. 125, 1390–1397, https://doi.org/10.1016/j.foodchem.2010.10.012

Theodoridou K., Yu P., 2013. Application potential of ATR-FT/IR molecular spectroscopy in animal nutrition: revelation of protein molecular structures of canola meal and presscake, as affected by heat-processing methods, in relationship with their protein digestive behavior and utilization for dairy cattle. J. Agric. Food Chem. 61, 5449–5458, https://doi.org/10.1021/jf400301y

Torres Mariani N.C., de Almeida Teixeira G.H., Gomes de Lima K.M., Morgenstern T.B., Nardini V., Cunha Jr L.C., 2015. NIRS and iSPA-PLS for predicting total anthocyanin content in jaboticaba fruit. Food Chem. 174, 643–648, https://doi.org/10.1016/j.foodchem.2014.11.008

Whitaker S., Pigford R.L., 1960. An approach to numerical differentiation of experimental data. Ind. Eng. Chem. 52, 185–187, https://doi.org/10.1021/ie50602a043

Williams P.C., Sobering D.C., 1993. Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. J. Near Infrared Sceptosc. 1, 25–32, https://doi.org/10.1255/jnirs.3

Wilson R.H., Belton P.S., 1988. A Fourier-transform infrared study of wheat starch gels. Carbohydr. Res.180, 339–344, https://doi.org/10.1016/0008-6215(88)80090-9

Windig W., Shaver J., Bro R., 2008. Loopy MSC: a simple way to improve multiplicative scatter correction. Appl. Spectrosc. 62, 1153–1159, https://doi.org/10.1366/000370208786049097

Wolkers W.F., Bochicchio A., Selvaggi G., Hoekstra F.A., 1998. Fourier transform infrared microspectroscopy detects changes in protein secondary structure associated with desiccation tolerance in developing maize embryos. Plant Physiol. 116, 1169–1177, https://doi.org/10.1104/pp.116.3.1169

Xu N., Liu J., Yu P., 2018. Using vibrational molecular spectroscopy to reveal association of steam-flaking induced carbohydrates molecular structural changes with grain fractionation, biodigestion and biodegradation. Spectrochim. Acta A Mol. Biomol. Spectrosc. 194, 181–188, https://doi.org/10.1016/j.saa.2018.01.019

Yang H., Irudayaraj J., Paradkar M.M., 2005. Discriminant analysis of edible oils and fats by FTIR, FT-NIR and FT-Raman spectroscopy. Food Chem. 93, 25–32, https://doi.org/10.1016/j.foodchem.2004.08.039

Yang J., Yen H.E., 2002. Early salt stress effects on the changes in chemical composition in leaves of iceplant and Arabidopsis. A Fourier transform infrared spectroscopy study. Plant Physiol. 130, 1032–1042, https://doi.org/10.1104/pp.004325

Zhang H., Ma J., Miao Y., Tuchiya T., Chen J.Y., 2015. Analysis of carbonyl value of frying oil by Fourier transform infrared spectroscopy. J. Oleo Sci. 64, 375–380, https://doi.org/10.5650/jos.ess14201