Submitted Paper

Applied Spectroscopy
www.s-a-s.org

Feasibility of In-Line Raman Spectroscopy for Quality Assessment in Food Industry: How Fast Can We Go?

Applied Spectroscopy
2022, Vol. 76(5) 559–568

DOI: 10.1177/00037028211056931
journals.sagepub.com/home/asp
SAGE

Tiril Aurora Lintvedt[1], Petter V. Andersen[1], Nils Kristian Afseth[1], Brian Marquardt[2], Lars Gidskehaug[3], and Jens Petter Wold[1]

Abstract

Raman spectroscopy is a viable tool within process analytical technologies due to recent technological advances. In this article, we evaluate the feasibility of Raman spectroscopy for in-line applications in the food industry by estimating the concentration of the fatty acids EPA + DHA in ground salmon samples (n = 63) and residual bone concentration in samples of mechanically recovered ground chicken (n = 66). The samples were measured under industry like conditions: They moved on a conveyor belt through a dark cabinet where they were scanned with a wide area illumination standoff Raman probe. Such a setup should be able to handle relevant industrial conveyor belt speeds, and it was studied how different speeds (i.e., exposure times) influenced the signal-to-noise ratio (SNR) of the Raman spectra as well as the corresponding model performance. For all samples we applied speeds that resulted in 1 s, 2 s, 4 s, and 10 s exposure times. Samples were scanned in both heterogenous and homogenous state. The slowest speed (10 s exposure) yielded prediction errors (RMSECV) of 0.41%EPA + DHA and 0.59% ash for the salmon and chicken data sets, respectively. The more in-line relevant exposure time of 1 s resulted in increased RMSECV values, 0.84% EPA + DHA and 0.84% ash, respectively. The increase in prediction error correlated closely with the decrease in SNR. Further improvements of model performance were possible through different noise reduction strategies. Model performance for homogenous and heterogenous samples was similar, suggesting that the presented Raman scanning approach has the potential to work well also on intact heterogenous foods. The estimation errors obtained at these high speeds are likely acceptable for industrial use, but successful strategies to increase SNR will be key for widespread in-line use in the food industry.

Keywords

Raman spectroscopy, process analytical technology, PAT, in-line food evaluation, representative sampling, omega-3 fatty acids, bone content

Date received: 20 July 2021; revised: 10 September 2021; accepted: 26 September 2022

Introduction

According to the United Nations, the percentage of food lost after harvesting and during transport, storage and processing is estimated to be 13.8% globally.[1] It is suggested that the production costs could be lowered and the efficiency of food systems increased if targeted interventions at critical stages of the value chain are implemented. Detailed in-line spectroscopic measurements of food raw materials can provide means for increased raw material utilization and ensuring stable product quality. Previously, several applications of near-infrared spectroscopy (NIRS) for evaluation of food products in the industry have been successfully realized, for example, rapid determination of edible meat content in crabs,[2] in-line fat distribution analysis of salmon fillets[3] and detection of woody breast syndrome in chicken fillets.[4] However, NIRS has limitations in chemical resolution because absorptions originate from overtones and combination modes which overlap intricately.[5] In situations where higher chemical specificity is needed, one option is to use Raman spectroscopy, which is based on

[1]Nofima AS, TromsØ, Norway
[2]MarqMetrix Inc, Seattle, WA, USA
[3]Aspen Technology Inc, Oslo, Norway

Corresponding author:
Tiril Aurora Lintvedt, Faculty of Science and Technology, NMBU, Nofima—Norwegian Institute for Food, Fisheries, and Aquaculture Research, Muninbakken 9-13, Breivika, Tromsø 9291, Norway.
Email: tiril.lintvedt@nofima.no

fundamental transitions and do not suffer from overlapping bands to the same extent. This can for instance be seen in spectra of fish oil as shown by Bekhit et al.[6] Raman spectroscopy is becoming a viable tool within process analytics due to recent technological advances. Low-cost Raman instruments are available and instrumental advances provide a new degree of versatility. The potential of in-line Raman applications have been demonstrated in several areas, such as in the pharmaceutical and bioprocessing domain, as reviewed by Esmonde-White et al.[7] However, the literature on Raman spectroscopy for in-line food evaluation is scarce, although the possibilities within the food industry are undoubtedly many. In addition, the existing studies on Raman-based strategies for reaction monitoring and control is often based on analysis of continuous and homogenous sample streams, while reported studies on in-line evaluation of single products moving along a conveyor belt is to a huge part lacking. The latter situation is frequently encountered in the food industry.

A prerequisite for in-line Raman applications is proper and representative optical sampling tools, and a development of particular interest in this respect is wide area illumination probes. This approach utilizes a defocused laser combined with multiple collection fibers, resulting in larger measurement areas and insensitivity to smaller variations in working distance.[8] This optical setup is suitable for measurements of food samples in a surface scanning mode. Recent work[9–12] demonstrates the potential for a surface scanning setup for food evaluation. Andersen et al. showed the applicability of Raman spectroscopy for bulk composition analysis of heterogenous foods. However, they used rather long exposure times ranging from 60 s to 80 s and pointed out that sampling speed could be a limitation for applications where single samples need to be analyzed in real-time. In other applications, Raman spectroscopy may give satisfactory measurements within milliseconds given the adequate laser power. In contrast, foods often have low Raman signals and may be damaged by exposure to higher laser powers, making in-line food applications more challenging.

One potential in-line application in the food industry is the evaluation of ground meat from the mechanical deboning process of rest-raw material of chicken, where remaining meat on the carcasses after filleting is separated from the bones. Perfect separation of bone and meat is not achieved, and the content of finely ground bone in the meat fraction is regulated.[13] In-line monitoring and control of the bone concentration is yet to be implemented, but Wubshet et al.[10] recently showed that Raman spectroscopy potentially could be used to quantify bone contents in these samples. Another relevant application is in-line determination of omega-3 fatty acids in salmon. Due to reported health effects, fatty acid composition is an important quality parameter in the market. NIRS has been successfully used for determination of total fat content in whole salmon fillets[14] and estimation of omega-3 fatty acids in pure fish oils is promising.[15,6] Brown et al.[16] reported that eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA) could be determined in salmon by NIR spectroscopy, but their calibration were less successful for intact salmon cutlets than for minced salmon and they suggest that the obtained performances are acceptable for rough approximation only. In addition, it is not clear whether the calibrations relied on measurements of the actual fatty acids or just a covariation with total fat as discussed by Eskildsen et al.[17] Raman spectroscopy is a promising tool for compositional analysis of fatty acids,[18–21] and in-line determination of omega-3 fatty acids in fillets can expand the opportunities for product differentiation. Furthermore, the implementation of such a spectroscopic sorting system on an industrial basis can provide the means for a rapid and affordable mapping of how feeding regimes affect the final fatty acid composition in the fish.

The main aim of this work was to evaluate the feasibility of Raman spectroscopy for in-line evaluation of complex foods, and to elucidate how fast a product can pass by on the conveyor belt while still obtaining a spectrum of sufficient signal-to-noise ratio (SNR) and acceptable modeling errors. This was investigated for sample sets based on (i) chicken rest-raw material from the mechanical deboning process and (ii) ground salmon belly trims. We used an industry relevant setup, applying the surface scanning strategy on samples moving along a conveyor belt. As food samples in general are heterogenous, we also aimed to confirm that the heterogeneity of the sample surface is not necessarily a limitation for a Raman scanning system. We did this by comparing prediction performance on heterogenous and homogenous versions of the same samples.

Material and methods

Raw Materials

Chicken Samples. Chicken samples came from a mechanical deboning process of rest-raw material of chicken. In this process, the remaining meat on the carcasses after filleting is separated from the bone fragments, resulting in two fractions: The mechanically deboned meat (MDM), containing mostly meat, and the mechanical deboning residues (MDR) containing mostly bone. Batches of MDM and MDR were provided by a poultry processing plant (Bioco, Nortura Hrland, Norway) and subsequently frozen. The batches were then ground in frozen state and used to make five different base blends with different bone concentrations. These blends were made by mixing ground MDM and MDR in different ratios, according to the following shares

$$\begin{aligned} &A: 100\%\text{MDR} \\ &B: 20\%\text{MDM} + 80\%\text{MDR} \\ &C: 50\%\text{MDM} + 50\%\text{MDR} \\ &D: 80\%\text{MDM} + 20\%\text{MDR} \\ &E: 100\%\text{MDM} \end{aligned}$$

Each sample was then made by combining four sub-samples of the available base blends A–E into the positions 1–4 in different arrangements in a rectangular sample holder (30 cm × 3 cm) as indicated in Fig. 1. This sample arrangement

was scanned with the Raman system. The aim of the arrangements was to create samples with different bone concentrations and varying heterogeneity. The bone content typically covaries with fat content in this kind of material. To reduce this correlation and to avoid overoptimistic modeling results, a fat-rich sub-sample (ground chicken skin) of varying size was added randomly to each sample (Fig. 1). A total of 66 different samples were made, spanning a realistic range of bone content. The samples were first measured as heterogenous compositions of base blends, then homogenized (Retsch Knife Mill Grindomix GM 200, 7000 rpm for 6 s twice with a stir in between) and measured again in the same sample holder. Five samples which consisted of only one base blend were not homogenized. The samples were made consecutively and the base blends were kept in a cold room (−1 °C–0 °C) throughout the whole experiment. Smaller portions of the blends were taken out to the experiment room (room temperature) at a time, to reduce possible temperature effects.

Salmon Samples. The samples were based on homogenized belly trims from 52 salmons acquired from three farming locations. Different feeding regimes were used on the different farming locations. Before homogenization (Retsch Knife Mill Grindomix GM 200, 7000 rpm for 3 s), the bellies were stored at (−1 °C–0 °C) to prevent liquid loss and lipid oxidation. The target for these samples was the concentration of omega-3 fatty acids EPA and DHA. To obtain an even distribution of fatty acid concentration in the sample set, 11 additional samples were made of 50/50 combination of two samples of the original set of 52. Those two sub-samples were placed side by side in the sample holder (Fig. 1). The total number of samples was then 63.

Reference Measurements

Reference analyses of the chicken samples were carried out by an external laboratory (ALS Laboratory Group, Oslo, Norway). Measurements for ash concentration (percent of wet weight) were carried out by gravimetric analysis (BS 4401 Part 1 1998 Commision Regulation (EC) 152/2009 MU 6,5%), and reference measurements for fat concentration (percent of wet weight) were carried out using pulsed nuclear magnetic resonance (NMR) analysis (MU 6.5%). Analyses of the salmon samples were carried out by BioLab (Bergen, Norway). Measurements of EPA and DHA concentration applied the reference standard AOCS Ce 1b-89 (methyl esterification capillary gas chromatography with a flame ionization detector or GC-FID), and was expressed as a percentage of the total amount of fatty acids (FA) in the analyzed sample. For all reference analyses the mean of two parallel sample measurements was reported.

Measurements and Data Analysis

Raman Measurements. For spectral acquisition, we employed a MarqMetrix all-in-one (AIO) Raman system equipped with a 785 nm laser operating at 450 mW power. The sampling optic used with the AIO was a wide area illumination (D = 3 mm), Proximal BallProbe HV standoff Raman probe with optimal working distance 8–10 cm (MarqMetrix Inc., Seattle, WA, USA). The probe was placed inside a tube structure in a dark cabinet on a conveyor belt, as shown in Fig. 2. The tube structure was perfectly fitted around the probe. The dark cabinet, designed for reducing ambient light, was a steel cabinet of 100 cm length and 5.2 cm height with black coating inside. For each measurement the Raman laser and signal acquisition were activated by a trigger system which relied on a laser sensor (CMOS laser sensor LR-ZB90CB from Keyence) for detection of a block placed beside the sample on the belt.

Each sample was placed on a plate covered with aluminum foil. The foil was used to prevent disturbing signals from the plastic conveyor belt upon potential imperfect sample triggering. Samples were passed through the dark cabinet and scanned at belt speeds 0.3 m/s, 0.15 m/s, 0.075 m/s, and 0.03 m/s. Corresponding exposure times were 1 s, 2 s, 4 s, and 10 s, respectively. The belt speed was tuned and controlled by using a contact tachometer (Tachometer PCE-DT 65 from PCE Instruments). Two technical replicates were measured for all samples and exposure times. Homogenized chicken samples were measured only for 2 s and 10 s exposures, for comparison with measurements on the corresponding heterogenous samples. Salmon samples were measured only in homogenized version, except for the 11 additional combination samples which were measured only in heterogenous state.

Pre-Processing.Chicken and salmon spectra were pre-processed using Savitzky–Golay (SG) smoothing (polynomial

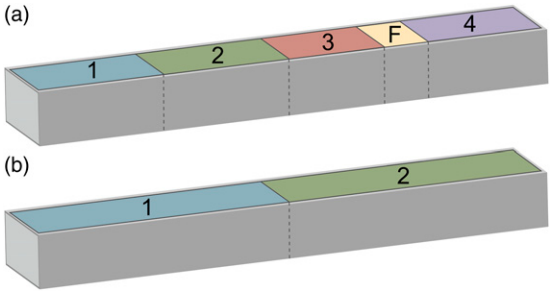

Figure 1. Sample composition scheme for chicken (a) and the 11 salmon combination samples (b). A fat-rich species F of varying size was placed at a random position in the chicken samples.

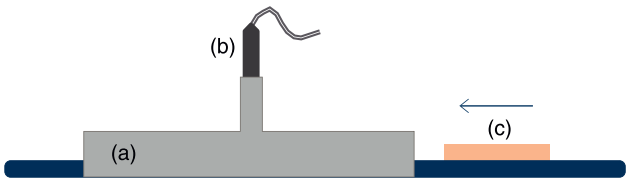

Figure 2. Spectrum acquisition setup consisting of a dark cabinet (a), wide area illumination standoff Raman probe (b) and a moving sample (c) on a conveyor belt.

order 2 and window size 9)[22] followed by extended multiplicative signal correction (EMSC)[23,24] employing up to the sixth order polynomial and the asymmetric least squares (ALS) algorithm[25,26] for baseline correction of the EMSC reference spectrum. The ALS reference spectrum correction employed a smoothing parameter of 5.8 (λ) and an asymmetric weighting parameter (of the residuals) of 0.01 (p). Pre-processing was applied on the full data sets, including all exposure times. Subsequently, the data sets were organized into separate exposure time subsets for data modeling.

Data Modeling. The Raman shift range 520 cm^{-1}–1800 cm^{-1} was utilized in the data modeling for both data sets. Partial least squares regression (PLSR)[27,28] was used for calibration development. We established models for ash concentration in chicken samples (proxy for bone content) and EPA + DHA concentration for salmon samples. Note that EPA and DHA concentrations were not estimated separately, but as a joint concentration value. The PLSR models were validated by cross-validation (CV) where replicate measurements were held out in the same segment to avoid overfitting. The reason for not averaging the replicates was to keep the prediction conditions as close to an in-line situation as possible. The choice of number of latent variables for the PLS modeling was based on a simple criterion using a 3% punish factor, as described by Westad and Martens.[29] Noise reduction and improvement in the model performance was attempted for the shorter exposure time subsets (4 s, 2 s, and 1 s) through variable selection employing the significance multivariate correlation (sMC) method.[30] For selection of variables in the 1 s exposure data, we applied sMC on the 2 s exposure data, and for selection of variables in the 2 s exposure data, sMC on the 4 s exposure data was used, and so on. This was to avoid overfitting. The results were compared with the corresponding calibrations based on the full spectrum (642 channels). Another method we investigated for noise reduction was averaging of in-line replicates. To evaluate this strategy, cross-validated PLSR models were obtained for three versions of the exposure time subsets; one version where the two spectrum replicates were averaged, one where replicate number 1 was selected as a representative for each sample and one where replicate number 2 was selected as a representative for each sample. We report the average performance of the two single-replicate versions as a representative for a single-measurement system.

The chicken samples were used to compare the performance of regression models on heterogenous samples versus homogenous samples. This comparison was conducted with CVANOVA,[31–33] a two-way analysis of variance (ANOVA) of cross-validation errors. In total five samples which consisted of only one base blend were excluded from this analysis, since these were not homogenized. The ANOVA was carried out in Python version 3.7, Anaconda3 distribution (Anaconda, Austin, TX). All other data analysis was carried out in Matlab version R2020a (The MathWorks, Natick, MA). The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Signal-to-Noise Ratio. Calculation of SNR was based on the ratio between the average spectrum intensity and the standard deviation of the estimated noise, similar to Guo et al.[34] An SNR value was calculated for each spectrum using Eq. 1.

$$\mathrm{SNR} = \frac{mean(I)}{sd(I_n)} \tag{1}$$

where I is the spectrum intensity and I_n is the estimated noise intensity. Noise was estimated as the difference between the spectrum and the smoothed version of the same spectrum, using SG with polynomial order 2 and window size 9. The SNR was compared across exposure times and data sets, and we have reported the average SNR of each exposure time subset (pre-processed versions).

Results and Discussion

Spectral Data

Figure 3 shows that there was a more prominent fluorescence background in the chicken spectra compared to the salmon spectra. This was expected since bone matrices are prone to fluoresce.[10] The fluorescence naturally increased with higher exposure times. From the pre-processed spectra (Fig. 4), the chemical bands are more easily distinguished. The chicken spectra consist mainly of bands associated with fatty acids (e.g., 1080, 1267, 1302, and 1441 cm^{-1}).[19,20] The peak at 1658 cm^{-1} may be associated with both the Amide I band related to proteins and the olefinic stretch related to fatty acids,[19,20] which overlap. More importantly, another prominent peak at 960 cm^{-1} is associated with phosphate ($\nu_1 PO_4^{3-}$), a well known bone mineral.[10,35] The salmon spectra are dominated by bands associated with fatty acids, similar to the chicken spectra. The most pronounced peaks are located in the region above 1200 cm^{-1}. This includes the bands at 1267 cm^{-1} and 1658 cm^{-1} which may be assigned to the Olefinic hydrogen bend and the Olefinic stretch, respectively, of which both can be related with unsaturated modes.[19,20] Peaks related to saturated modes, that is, at 1302 cm^{-1} and 1441 cm^{-1}, can be assigned to the methylene twisting deformations and the Methylene scissor deformations, respectively.[19,20] In the region below 1200 cm^{-1}, two noticeable peaks are located at 1004 cm^{-1} and 1080 cm^{-1}, which originate from the aromatic ring breathing of phenylalanine, and the liquid aliphatic C–C stretch in gauche, respectively.[19,20] Another important peak at 935 cm^{-1} is most likely related to the alkene C–H deformation in polyunsaturated fatty acid moieties.[36,37]

Implication of Conveyor Belt Speed

Figure 4 shows pre-processed spectra at different exposure times for both sample sets. As indicated, the SNR clearly decreased with higher scanning speed. Note also that the

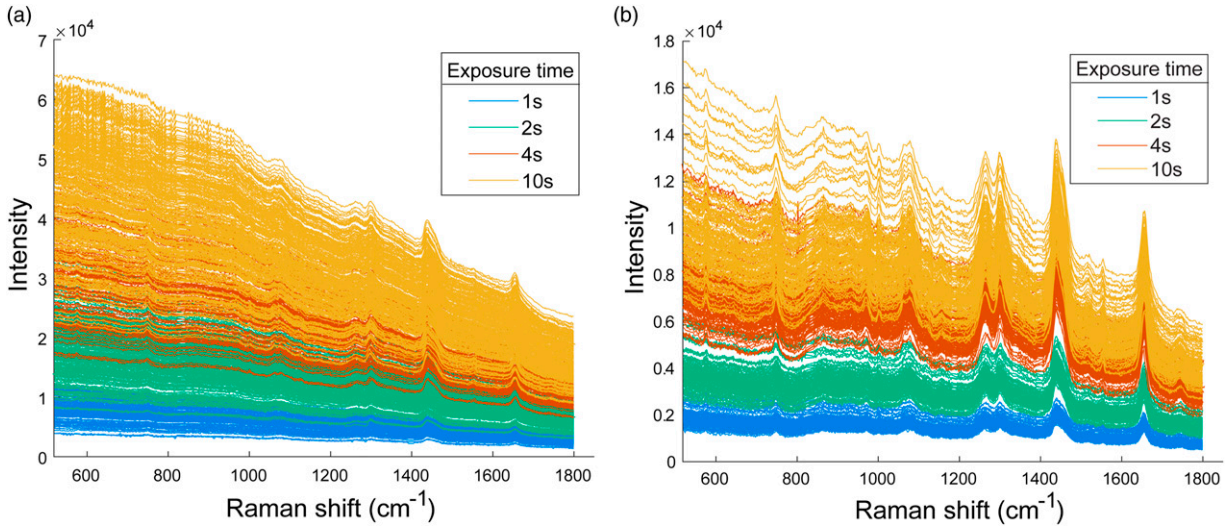

Figure 3. Raw spectra from chicken (a) and salmon samples (b), colored according to exposure time.

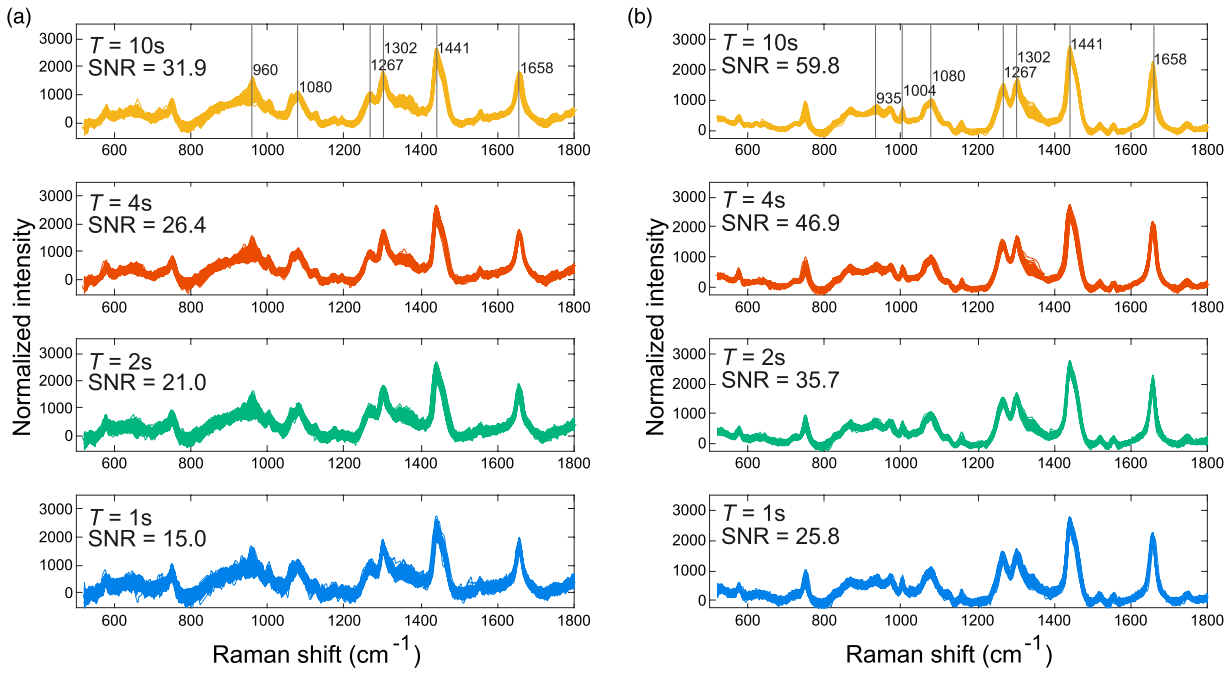

Figure 4. Pre-processed sample spectra from chicken (a) and salmon (b) for all exposure times *T*.

spectra from chicken had considerably lower SNR than those from salmon. Hence, in this context the chicken measurements represent low SNR spectra, while the salmon measurements represent high SNR spectra. The regression results for both sample sets and all exposure times are summarized in Fig. 5 (see additional results in the Supplemental Material). There was a clear trend in performance as function of exposure time for both sample sets, and as expected we obtained higher root mean square errors (RMSECV) for shorter exposure times. This was in essence due to lower SNR, as apparent from the strong negative correlations between these two ($r = -0.96$ for chicken and $r = -0.99$ for salmon). Nevertheless, the models based on different exposure times overall showed the same chemical signatures in the regression vectors (Fig. 6), which was encouraging. Note, however, that as the exposure time decreased, the most pronounced regions in the regression vectors became less weighted.

The exposure time of 10 s was regarded a best case reference and represented a rather slow conveyor belt speed, while the 1–2 s exposure times represented highly relevant conveyor belt speeds. With respect to model performance, there seemed to be a critical exposure time limit around

3 s–4 s. The graphs (Fig. 5) indicate that higher exposure times gave only a marginal increase in performance, while decreasing the exposure time below this limit had a more detrimental effect. Such critical limits will vary depending on the food properties related to Raman scattering cross sections and for different target compounds. For instance, measurements on low fat salmon muscle might increase the critical exposure time limit due to weaker signals and lower SNR. Although the estimation errors obtained at the lower exposure times in these particular cases are close to acceptable for industrial use, the rapid decrease in performance below 3 s–4 s illustrates that it would be a definitive advantage with developments on instrument sensitivity and other efforts on SNR optimization. In spite of the overall higher SNR levels and the same relative decrease in SNR from 4 s to 2 s, there was a more dramatic effect on the model performance for salmon than chicken (Fig. 5). This may be because the %ash model relies mainly on a single peak while the %EPA + DHA model relies on more subtle spectroscopic changes in several peaks, as can be seen from the regression coefficients in Fig. 6. This emphasizes that the critical SNR level in a predictive might system depend on the complexity of the model.

In the closely related work by Wubshet et al.,[10] where a Raman system (standoff probe, 785 nm, spot size 6 mm) was utilized for %ash estimation in similar chicken samples, an RMSECV of 0.63% (of wet weight) was obtained with an accumulation time of 15 s × 4. We obtained a corresponding RMSECV with 4 s exposure, which emphasize that the sampling strategy employed in this study performs very well. Considering that the exposure time was significantly decreased from 10 s to 1 s, the resulting error increase of only 0.2% is promising, also taken into account that the error in the reference measurements of ash concentration was 6.5% of measured value, similar to the modeling error. The reference uncertainty gives a limit for how well our model can perform. However, one should consider in more detail the implication of the exposure time reduction on the model itself. The main observation (Fig. 5) was a decrease in number of latent variables (LVs) going from 2 s to 1 s exposure time, indicating that we lose some information by employing an exposure time of 1 s. The regression vectors in Fig. 6 clearly show that predictions are mainly based on the mineral band at 960 cm^{-1}. Moreover, it can be noted that fat associated peaks are not prominent in the chicken model regression vectors, and the correlation between %fat and %ash reference values (r =−0.65) is not critical.

The regression coefficients for the %EPA + DHA models (Fig. 6) show that the main positively correlated peaks are at 935 cm^{-1}, 1264 cm^{-1}, and 1663 cm^{-1}. The first is most likely related to the alkene C–H deformation in polyunsaturated fatty acid moieties.[36,37] The two latter can be associated with other unsaturated modes.[20,19] The main negatively correlated peaks are at 1004 cm^{-1}, 1081 cm^{-1}, 1305 cm^{-1}, and 1443 cm^{-1}, of which the two latter can both be associated with saturated fatty acids.[20,19] In such PLSR models there will always be some uncertainty with respect to indirect modeling on other constituents which co-vary with the analyte, as discussed thoroughly by Eskildsen et al.[17] In this sample set, reference values for %EPA + DHA did not correlate strongly with total polyunsaturated fatty acids (r = 0.40), but correlations were more evident with total monounsaturated fatty acids (r = −0.77) and total saturated fatty acids (r = 0.88). However, peaks associated with unsaturated modes are clearly important for the model. All peaks mentioned above are visible in models based on different exposure times. However, the more subtle details visible in the 10 s exposure model are gradually more compromised with decreasing exposure time. Additionally, the models of exposure times below 4s have a decreasing complexity with respect to number of LVs (Fig. 5). This reinforces the impression that we lose information by decreasing exposure time below the critical exposure time around 3 s–4 s.

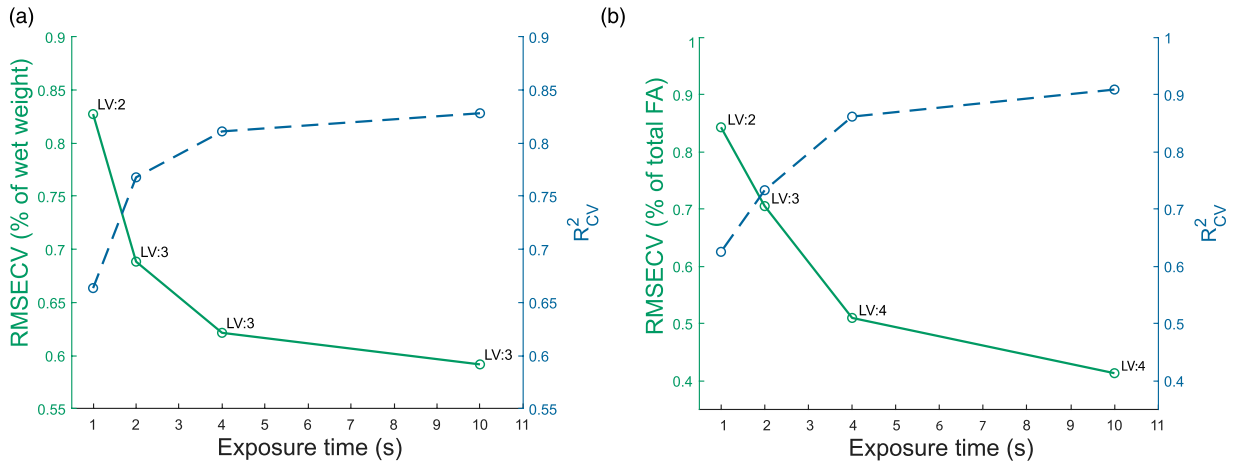

Figure 5. Model performance metrics across different exposure times for ash estimation in chicken (a) and EPA+DHA estimation in salmon (b). We show the RMSECV (solid line) and the coefficient of determination (R^2_{CV} in dashed line). The number of latent variables (LV) employed for each calibration is indicated.

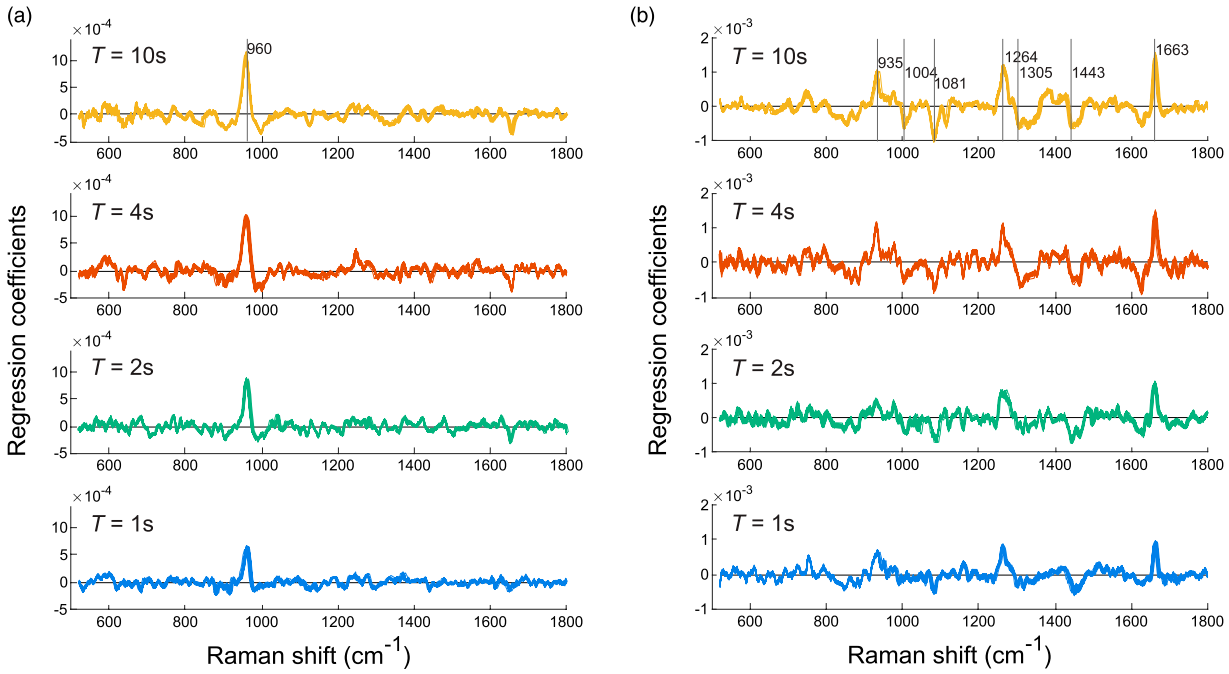

Figure 6. Regression vectors for PLSR models for ash in chicken (a) and EPA+DHA in salmon (b) for all exposure times *T*.

Reduction of Noise

Since SNR is the main limiting factor for model performance, three different strategies for noise reduction were investigated for potential improvement. Variable selection was motivated by the apparently uninformative regions in the regression vectors (Fig. 6) which could possibly contribute with noise. Cross-validated results with applied sMC variable selection are summarized in Table I. The effect on the performance was moderate for 4 s and 2 s exposure times for both chicken and salmon, but resulted in a reduction in number of LVs in the models, which indicate more robust models. For 1 s exposure time, the effect of variable selection on the performance was slightly more pronounced in both chicken and salmon data.

Another possible strategy to increase SNR is to place two instruments in series on the conveyor belt and make predictions based on average spectra. In Table II we compare the average performance based on the two single-replicate versions of the data sets to the replicate-average versions. For both chicken and salmon, the effect of replicate averaging was evident. For chicken, the effect was particularly apparent for 1s exposure time, where the RMSECV decreased by 0.18 %ash. For salmon, the effect was more evenly pronounced across exposure times, with an RMSECV decrease around 0.1 %EPA + DHA for the 1 s–4 s exposures, while the impact for the 10 s exposure was less discernible. Note that in most cases, the effect of replicate averaging was approximately the same as doubling the exposure time for the single-replicate set.

Table I. PLSR results for ash in chicken (% of wet weight) and EPA+DHA in salmon (% of total FA) before and after variable selection.

	Full spectrum		Selected variables		
Exp. time	RMSECV	LV[a]	RMSECV	LV	No. variables[b]
Chicken					
1 s	0.84	2	0.76	2	199
2 s	0.70	3	0.68	2	204
4 s	0.63	3	0.61	2	211
Salmon					
1 s	0.84	2	0.78	2	217
2 s	0.71	3	0.70	2	192
4 s	0.51	4	0.53	2	217

[a]Latent variables.
[b]Number of selected variables.

From another practical perspective, products on the conveyor belt may vary in size and it might be desirable to adjust exposure time according to product size in order to optimize the SNR for each product. To investigate if measurements from different exposure times can be used in the same calibration, we applied the 10 s exposure model on the shorter exposure data (4 s, 2 s, and 1 s) in a cross-validation scheme. Two replicates were held out of all exposure time subsets for each validation segment, where a model was built on the 10 s exposure data and applied on the held out samples of the shorter exposure data. The number of LVs included in the models were predefined as three for %ash

and four for %EPA + DHA estimation. Results are summarized in Table III. The two data sets showed very similar results when basing predictions on the 10 s exposure model as when applying the original cross-validation scheme for the separate exposure time subsets. Thus, it indicates that a model built on 10 s exposure time measurements works just as well for the spectra of 1 s exposure. This demonstrates that smaller variations in exposure time between samples are not critical to the performance of the system. A pre-processing step involving spectral normalization is essential for this approach to work. One practical instrument issue when varying exposure times may be the occurrence of mismatch between acquired spectra and dark spectra. Strategies to tackle this, either through pre-processing or practical solutions, should be considered.

Table II. PLSR results for ash in chicken (% of wet weight) and EPA+DHA in salmon (% of total FA), using (i) only one of the replicates for each sample and (ii) the mean of two replicates.

	One selected replicate		Replicate average	
Exp. time	$RMSECV_{avg}$	LV[a]	RMSECV	LV
Chicken				
1 s	0.88	2	0.70	3
2 s	0.68	3	0.62	3
4 s	0.62	3	0.53	3
10 s	0.58	3	0.51	3
Salmon				
1 s	0.85	2,3	0.75	2
2 s	0.74	3	0.64	3
4 s	0.54	3,4	0.43	4
10 s	0.43	4,3	0.38	4

[a]Latent variables.

Table III. PLSR results for ash in chicken (% of wet weight) and EPA+DHA in salmon (% of total FA), using a calibration based on 10 s exposure spectra on the shorter exposure data sets. Results from the original cross-validation scheme is included.

		10 s calibration	
Exp. time	Original RMSECV	RMSECV	LV[a]
Chicken			
1 s	0.84	0.83	3
2 s	0.70	0.70	3
4 s	0.63	0.62	3
Salmon			
1 s	0.84	0.82	4
2 s	0.71	0.74	4
4 s	0.51	0.49	4

[a]Latent variables.

Impact of Heterogeneity on the Raman Scanning Measurements

The results from CVANOVA showed that the differences in cross-validation errors for homogenized and heterogenous chicken samples were not significantly different for either the 2 s or 10 s exposure times (p-values of 0.47 and 0.36, respectively). This indicates that the heterogeneity itself is not a major challenge for the scanning strategy. However, it should be noted that the sample surfaces in this experiment were representative for the sample bulk composition and that the heterogeneity of certain intact foods such as fish fillets is more complex. Due to the still limited sampling volume of a Raman scan, it is important to consider appropriate sampling strategies in more detail when we encounter heterogeneity in several dimensions of a product (e.g., in depth and transverse to the scanning direction). An important step for suggesting a robust sampling regime in such situations is to map the composition profile of the product. In this work, we used chicken samples with carefully designed heterogeneity and salmon samples which were homogenous. Nevertheless, the results provide a useful insight towards determining the feasibility of a Raman surface scanning strategy, also for intact and more heterogenous products like salmon fillets.

Conclusion

We have shown that spectra obtained from in-line Raman scanning of single chicken and salmon samples have sufficiently high quality for exposure times ranging from 10 s to 1 s. With appropriate strategy developments, it is viable to use a wide area illumination standoff Raman probe for fast in-line evaluation of % ash and %EPA + DHA in complex foods. The SNR clearly decreases with higher scanning speeds, and it was evident from model performances that SNR is a critical parameter. Efforts to optimize SNR in a given system is therefore important. This can be achieved through instrument improvements, variable selection and also by strategies based on flexible exposure times or using Raman instruments in series. Furthermore, we confirmed that the heterogeneity of the sample surface is not necessarily a limitation for the scanning strategy, but for food products where heterogeneity is more complex, it is likely important to consider individual sampling strategies.

Acknowledgments

We would like to express our gratitude to Katinka Dankel and Karen Wahlstrøm Sanden at Nofima for assistance during sample preparation and data acquisition.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was partially funded by the Research Council of Norway through the projects SFI Digital Food Quality [grant number 309259] and the Food Pilot Plant [grant number 296083], along with the Norwegian Agricultural Food Research Foundation through the project Precision Food Production [grant number 314111].

ORCID iDs

Tiril Aurora Lintvedt https://orcid.org/0000-0001-8527-0138
Petter V. Andersen https://orcid.org/0000-0002-8241-1051

Supplemental Material

The supplemental material mentioned in the text, consisting of Figures S1, is available in the online version of the journal

References

1. United Nations Department of Economic and Social Affairs. "Responsible consumption and production". In: The Sustainable Development Goals Report. 2020. Pp. 48–49. doi: 10.18356/214e6642-en.
2. J.P. Wold, M. Kermit, A. Woll. "Rapid Nondestructive Determination of Edible Meat Content in Crabs (*Cancer Pagurus*) by Near-Infrared Imaging Spectroscopy". Appl. Spectrosc. 2010. 64(7): 691–699. doi: 10.1366/000370210791666273.
3. V.H. Segtnan, M. Høy, F. Lundby, B. Narum, J.P. Wold. "Fat Distribution Analysis in Salmon Fillets Using Non-Contact Near Infrared Interactance Imaging: A Sampling and Calibration Strategy". J. Near Infrared Spectrosc. 2009. 17(5): 247–253. doi: 10.1255/jnirs.851.
4. J.P. Wold, I. Måge, A. Løvland, K.W. Sanden, R. Ofstad. "Near-infrared Spectroscopy Detects Woody Breast Syndrome in Chicken Fillets by the Markers Protein Content and Degree of Water Binding". Poult. Sci. 2019. 98(1): 480–490. doi: 10.3382/ps/pey351.
5. Y. Ozaki, C.W. Huck, K.B. Beć. "Near-IR Spectroscopy and its Applications". In: V.P. Gupta, editor. Molecular and Laser Spectroscopy: Advances and Applications. Amsterdam, Netherlands: Elsevier, 2017. Pp 11–38.
6. M.Y. Bekhit, B. Grung, S.A. Mjøs. "Determination of Omega-3 Fatty Acids in Fish Oil Supplements Using Vibrational Spectroscopy and Chemometric Methods". Appl. Spectrosc. 2014. 68(10): 1190–1200. doi: 10.1366/13-07210.
7. K.A. Esmonde-White, M. Cuellar, C. Uerpmann, L. Bruno, I.R. Lewis. "Raman Spectroscopy as a Process Analytical Technology for Pharmaceutical Manufacturing and Bioprocessing". Anal. Bioanal. Chem. 2017. 409: 637–649. doi:10.1007/s00216-016-9824-1.
8. H. Wikström, I.R. Lewis, L.S. Taylor. "Comparison of Sampling Techniques for In-Line Monitoring Using Raman Spectroscopy". Appl. Spectrosc. 2005. 59(7): 934–941. doi: 10.1366/0003702054411553.
9. S.G. Wubshet, J.P. Wold, N.K. Afseth, U. Böcker, D. Lindberg. "Feed-Forward Prediction of Product Qualities in Enzymatic Protein Hydrolysis of Poultry By-products: A Spectroscopic Approach". Food Bioprocess. Technol. 2018. 11: 2032–2043. doi: 10.1007/s11947-018-2161-y.
10. S.G. Wubshet, J.P. Wold, U. Böcker. "Raman Spectroscopy for Quantification of Residual Calcium and Total Ash in Mechanically Deboned Chicken Meat". Food Control. 2019. 95: 267–273. doi: 10.1016/j.foodcont.2018.08.017.
11. O. Monago-Maraña, J.P. Wold, R. Rødbotten, K.R. Dankel, N.K. Afseth. "Raman, Near-infrared and Fluorescence Spectroscopy for Determination of Collagen Content in Ground Meat and Poultry By-products". Food Sci. Technol. 2021. 140: 110592. doi: 10.1016/j.lwt.2020.110592.
12. P.V. Andersen, J.P. Wold, N.K. Afseth. "Assessment of Bulk Composition of Heterogeneous Food Matrices Using Raman Spectroscopy". Appl. Spectrosc. 2021. 75(10): 1278–1287. doi: 10.1177/0003702821006150.
13. EFSA. "Scientific Opinion on the Public Health Risks Related to Mechanically Separated Meat (MSM) Derived from Poultry and Swine". EFSA J. 2013. 11(3): 3137. doi: 10.2903/j.efsa.2013.3137.
14. J.P. Wold, M. Kermit, V.H. Segtnan. "Chemical Imaging of Heterogeneous Muscle Foods Using Near-Infrared Hyperspectral Imaging in Transmission Mode". Appl. Spectrosc. 2016. 70(6): 953–961. doi: 10.1177/0003702816641260.
15. M. M Cascant, C. Breil, A.S. Fabiano-Tixier, F. Chemat, S. Garrigues. "Determination of Fatty Acids and Lipid Classes in Salmon Oil by Near Infrared Spectroscopy". Food Chem. 2018. 239: 865–871. doi: 10.1016/j.foodchem.2017.06.158.
16. M.R. Brown, P.D. Kube, R.S. Taylor, N.G. Elliott. "Rapid Compositional Analysis of Atlantic Salmon (*Salmo Salar*) Using Visible-Near Infrared Reflectance Spectroscopy". Aquacult. Res. 2014. 45(5): 798–811. doi: 10.1111/are.12021.
17. C.E. Eskildsen, T. Næs, P.B. Skou, L.E. Solberg, K.R. Dankel. "Cage of Covariance in Calibration Modeling: Regressing Multiple and Strongly Correlated Response Variables onto a Low Rank Subspace of Explanatory Variables". Chemom. Intell. Lab. Syst. 2021. 213: 104311. doi: 10.1016/j.chemolab.2021.104311.
18. N.K. Afseth, V.H. Segtnan, B.J. Marquardt, J.P. Wold. "Raman and Near-Infrared Spectroscopy for Quantification of Fat Composition in a Complex Food Model System". Appl. Spectrosc. 2005. 59(11): 1324–1332. doi: 10.1366/000370205774783304.
19. N.K. Afseth, J.P. Wold, V.H. Segtnan. "The Potential of Raman Spectroscopy for Characterisation of the Fatty Acid Unsaturation of Salmon". Anal. Chim. Acta. 2006. 572(1): 85–92. doi: 10.1016/j.aca.2006.05.013.
20. J.R. Beattie, S.E. Bell, C. Borgaard, A. Fearon, B.W. Moss. "Prediction of Adipose Tissue Composition Using Raman Spectroscopy: Average Properties and Individual Fatty Acids". Lipids. 2006. 41(3): 287–294. doi: 10.1007/s11745-006-5099-1.

21. D.P. Killeen, S.N. Marshall, E.J. Burgess, K.C. Gordon, N.B. Perry. "Raman Spectroscopy of Fish Oil Capsules: Polyunsaturated Fatty Acid Quantitation Plus Detection of Ethyl Esters and Oxidation". J. Agric. Food Chem. 2017. 65(17): 3551–3558. doi: 10.1021/acs.jafc.7b00099.
22. A. Savitzky, M.J. Golay. "Smoothing and Differentiation of Data by Simplified Least Squares Procedures". Anal. Chem. 1964. 36(8): 1627–1639. doi: 10.1021/ac60214a047.
23. H. Martens, E. Stark. "Extended Multiplicative Signal Correction and Spectral Interference Subtraction: New Preprocessing Methods for Near Infrared Spectroscopy". J. Pharm. Biomed. Anal. 1991. 9(8): 625–635. doi: 10.1016/0731-7085(91)80188-F.
24. K.H. Liland, A. Kohler, N.K. Afseth. "Model-Based Pre-Processing in Raman Spectroscopy of Biological Samples". J. Raman Spectrosc. 2016. 47(6): 643–650. doi: 10.1002/jrs.4886.
25. P.H.C. Eilers, H.F.M. Boelens. "A Perfect Smoother". Life Sci. 2003. 75(14): 3631–3636. doi: 10.1021/ac034173t.
26. K.H. Liland, T. Almøy, B.H. Mevik. "Optimal Choice of Baseline Correction for Multivariate Calibration of Spectra". Appl. Spectrosc. 2010. 64(9): 1007–1016. doi: 10.1366/000370210792434350.
27. H. Martens, T. Næs. Multivariate Calibration. Chichester, UK: John Wiley and Sons, 1989. Pp. 116–165.
28. Å. Björck, U.G. Indahl. "Fast and Stable Partial Least Squares Modelling: A Benchmark Study with Theoretical Comments". J. Chemom. 2017. 31(8). doi: 10.1002/cem.2898.
29. F. Westad, H. Martens. "Variable Selection in Near Infrared Spectroscopy Based on Significance Testing in Partial Least Squares Regression". J. Near Infrared Spectrosc. 2000. 8(2): 117–124. doi: 10.1255/jnirs.271.
30. T.N. Tran, N.L. Afanador, L.M. Buydens, L. Blanchet. "Interpretation of Variable Importance in Partial Least Squares with Significance Multivariate Correlation (sMC)". Chemom. Intell. Lab. Syst. 2014. 138: 153–160. doi: 10.1016/j.chemolab.2014.08.005.
31. L. Eriksson, J. Trygg, S. Wold. "CV-ANOVA for Significance Testing of PLS and OPLS Models". J. Chemom. 2008. 22(11–12): 594–600. doi: 10.1002/cem.1187.
32. U.G. Indahl, T. Næs. "Evaluation of Alternative Spectral Feature Extraction Methods of Textural Images for Multivariate Modeling". J. Chemom. 1998. 12(4): 261–278. doi: 10.1002/(sici)1099-128x(199807/08)12:4{\textless}261::aid-cem513{\textgreater}3.3.co;2-q.
33. J.P. Wold, K. Kvaal, B. Egelandsdal. "Quantification of Intramuscular Fat Content in Beef by Combining Autofluorescence Spectra and Autofluorescence Images". Appl. Spectrosc. 1999. 53(4): 448–456. doi: 10.1366/00037029991946730.
34. S. Guo, C. Beleites, U. Neugebauer, S. Abalde-Cela, N.K. Afseth. "Comparability of Raman Spectroscopic Configurations: A Large Scale Cross-Laboratory Study". Anal. Chem. 2020. 92(24): 15745–15756. doi: 10.1021/acs.analchem.0c02696.
35. M.D. Morris, G.S. Mandair. "Raman Assessment of Bone Quality". Clin. Orthop. Relat. Res. 2011. 469(8): 2160–2169. doi: 10.1007/s11999-010-1692-y.
36. K. Czamara, K. Majzner, M.Z. Pacia, K. Kochan, A. Kaczor. "Raman Spectroscopy of Lipids: A Review". J. Raman Spectrosc. 2015. 46(1): 4–20. doi: 10.1002/jrs.4607.
37. G. Socrates. "Alkenes, Oximes, Imines, Amidines, Azo compounds: C=C, C=N, N=N Groups". In: G. Socrates, editor. Infrared and Raman Characteristic Group Frequencies: Tables and Charts. Chichester, UK: John Wiley and Sons, 2004. p. 74.