Assessment of Bulk Composition of Heterogeneous Food Matrices Using Raman Spectroscopy

Petter Vejle Andersen [ID]**, Jens Petter Wold, and Nils Kristian Afseth**

Abstract

Raman spectroscopy (RS) has for decades been considered a promising tool for food analysis, but widespread adoption has been held back by, e.g., high instrument costs and sampling limitations regarding heterogeneous samples. The aim of the present study was to use wide area RS in conjunction with surface scanning to overcome the obstacle of heterogeneity. Four different food matrices were scanned (intact and homogenized pork and by-products from salmon and poultry processing) and the bulk chemical parameters such as fat and protein content were estimated using partial least squares regression (PLSR). The performance of PLSR models from RS was compared with near-infrared spectroscopy (NIRS). Good to excellent results were obtained with PLSR models from RS for estimation of fat content in all food matrices (coefficient of determination for cross-validation (R^2_{CV}) from 0.73 to 0.96 and root mean square error of cross-validation (RMSECV) from 0.43% to 2.06%). Poor to very good PLSR models were obtained for estimation of protein content in salmon and poultry by-product using RS (R^2_{CV} from 0.56 to 0.92 and RMSECV from 0.85% to 0.94%). The performance of RS was similar to NIRS for all analyses. This work demonstrates the applicability of RS to analyze bulk composition in heterogeneous food matrices and paves way for future applications of RS in routine food analyses.

Introduction

For several decades, Raman spectroscopy (RS) has been considered a potential and promising tool for food analysis.[1,2] In the last 25 years, the field of RS has evolved from academic laboratory instrumentation to robust commercially available solution-based systems.[3] But still, a very limited number of RS applications has found its way into commercial use in the food industry. Since food matrices often are heterogeneous samples of high chemical complexity, a major requisite for reliable food analysis is related to representative sampling. The lack of appropriate tools for representative sampling can therefore be highlighted as one of the several factors that has prevented a widespread use of RS in food analysis.

The unique potential of RS in food analysis is related to the ability to capture subtle chemical distinctions in food matrices in a nondestructive way. Whereas techniques such as near-infrared spectroscopy (NIRS) provide commercial solutions for assessment of bulk parameters such as fat and protein contents, RS can potentially provide more accurate estimates of features such as fatty acid composition,[4]

protein composition,[5] and pigmentation.[6] To benefit from the latter possibilities, it would be advantageous if RS could be used also to assess bulk chemical properties. In this way, potentially expensive and complex hyphenated systems (e.g., combinations of NIRS and RS) to cover a range of different analytes would be avoided. An example of a bulk chemical parameter in foods is intramuscular fat (IMF) content in meat. Since the fat is dispersed unevenly within the muscle, it stresses the need for sampling a relatively large volume of the muscle to get a representative spectrum. Fowler et al.[7] used a handheld Raman system with a spot size diameter of $50 \, \mu m$ to assess IMF in lamb muscle, but results were not encouraging, coefficient of determination (R^2_{CV}) $= 0.02$ and root mean square error of cross-validation (RMSECV) $= 1.12\%$, indicating that sampling procedures had to be improved. Collecting Raman spectra from

Nofima, Ås, Norway

Corresponding author:
Petter Vejle Andersen, Nofima, Osloveien 1, 1431 Ås, Norway.
Email: petter.andersen@nofima.no

homogenized meat,[8] intact meat sampled at many spots,[9] or using a wide area probe on three spots[10] have shown improvements (R^2_{CV} from 0.64 to 0.85 with corresponding RMSECV from 0.80% to 0.09%) but results are still not conclusive. This emphasizes the need for further development of representative sampling tools for RS.

Traditionally, Raman instrumentation has been restricted to very small sample spots (often below a diameter of 1 mm). In addition, the conventional Raman backscattering geometry has confined Raman analysis to the near surfaces of a sample.[11] Spatially offset Raman spectroscopy (SORS) and transmission RS are alternative sampling approaches providing means for depth penetration and representative sampling.[12,13] SORS is a relevant tool for food analysis and is especially suited for layered food samples, such as for instance intact fish, where analysis is dependent on penetrating the skin.[14,15] Transmission RS is highly feasible in food applications of samples of adequate thickness (typically less than 20 mm), and the technique has been used to assess bulk properties in different food matrices such as protein contents of packed corn kernels[16] and protein and oil contents in soy beans.[17,18] However, for food matrices of greater and varying thickness, such as typical muscle food samples, a backscattering geometry is a preferable approach.

The shortcomings of the conventional backscattering geometry related to representative sampling have been met by different strategies. Sample rotation during spectral collection is among the more advanced approaches. An example of this approach was provided by Beattie et al.,[19] who used a laser, focused as a line 7 mm long, to acquire Raman spectra of roasted meat tissues clipped to a rotating stage. In this way, an estimated sampling volume of 1.5 cm³ was obtained. One of the milestones in representative Raman analysis has been the development of wide area RS, also denoted volumetric or global illumination Raman. Here, a defocused laser beam is used to illuminate a larger area within the sample, and multiple collection fibers covering the same sample area are used for photon collection.[3,20,21] Wikström et al.[22] used this approach in the comparison of sampling techniques for in-line monitoring of pharmaceutical unit operations (i.e., wet granulation process). Their results revealed that wide area RS provided significant advantages over standard backscattering approaches both in terms of enhanced sampling volumes and robustness towards fluctuations in sampling distances. Subsequently, commercial solutions for wide area RS have been made available, and the technique has found a range of applications within pharmaceutical, biomedical, and chemical analysis.[11] Surprisingly, wide area RS has not been much used in food analysis. Kim et al.[23] have shown that wide area RS improves classification of rice into geographic origins compared to the conventional backscattering geometry. We have used a commercial wide area RS system to explore Raman analysis in a variety of applications, from

prediction of water holding capacity and pH in porcine meat[10,24] to measuring ASTA color values (American Spice Trade Association standard for measuring color in spices) and Sudan I content in paprika powder.[25] Recently, it was shown that wide area RS could be used in scanning mode, by moving the laser probe over the sample during analysis. With such an approach, Wubshet et al.[26] showed the first application of RS as a rapid tool for estimating calcium and ash contents in bone and meat mixtures after mechanical deboning of chicken meat. The same approach was recently also used to characterize chemical variation of poultry by-products fed into a bioprocess[27] and for quantification of collagen contents in ground meat.[5]

Moving a wide area RS probe around a sample surface opens the possibility for representative sampling of rather heterogeneous samples. Therefore, in the present study, wide area RS was evaluated for assessment of bulk chemical properties (i.e., fat and protein contents) of heterogeneous food matrices. Three different sample sets comprising three different food matrices, and thus different levels of heterogeneity, were studied: (i) fine ground poultry by-products; (ii) fresh intact and homogenized pork; and (iii) coarsely ground salmon by-products. For all sample sets, the performance of RS was compared to results obtained using NIRS. To the authors' knowledge, this is the first time wide area RS is critically evaluated for assessing bulk chemical properties of foods.

Materials and Methods

Sampling of Pork, Poultry, and Salmon By-Products

A 5 cm thick slice of longissimus lumborum was excised from pig carcasses ($n = 99$) at the cutting line two days post-mortem. All animals were slaughtered in accordance with Norwegian guidelines in a commercial slaughterhouse. Sample collection was carried out over two two-day periods, with two weeks in between each sampling. Sampling was not controlled for age, sex, or breed of the pigs. A subsample from the center of longissimus lumborum sample weighing approx. 80 g was cut out, vacuum packed, and stored at −20 °C. Samples were thawed at 4 °C overnight and spectra from intact samples were recorded with Raman and NIRS. After spectra were recorded from the intact samples, each sample was cut into smaller pieces and homogenized for approximately 5 s using a Krups Type 708 A food processor (Krups, Germany). Homogenized samples were scanned again with Raman and NIRS. Samples were subsequently frozen at −20 °C for determination of IMF content.

The poultry by-product sample material was collected from a poultry processing plant (Bioco, Norway). Five by-product fractions were selected, consisting of chicken skin, and carcasses from chicken and turkey, both before and after mechanical deboning. These fractions were either

pure or mixed together to make 52 samples with unique chemical composition. The samples were finely ground on-site in an industrial grinder (Seydelmann, Germany) with a 7.9 mm grinder plate and immediately measured with NIRS. A sub-sample of 400 g was taken out and stored at 4 °C until further analysis by RS.

Fifty-five samples of approximately 3 kg of by-products from the production of Atlantic salmon (*Salmo salar*) were obtained from a Norwegian by-product processing plant (Biomega AS, Norway). The samples included varying amounts of heads, skin, backbones (with muscle residues), gutted whole salmon, and belly flaps/trimmings, in addition to mixtures of these by-products. The samples were coarsely ground in an industrial Wolfking grinder (Wolfking Inc., USA) with a 30 mm grinder plate prior to the NIRS and Raman measurements. After spectroscopic investigations, the ground samples were further homogenized in a Grindomix GM 200 (Retsch, Germany) before fat and protein content determination.

Raman Spectroscopy

Raman spectra were recorded with a Kaiser RamanRXN2 multichannel Raman analyzer (Kaiser Optical Systems, Inc., USA) equipped with a 785 nm laser at 400 mW and a probe hybridization array typing probe operating at a working distance of 250 mm. At optimal working distance, the laser spot size is 6 mm in diameter on the sample. The collection fibers of the probe are arranged so that Raman scattering is collected from a wider area than the point of illumination. With current excitation wavelength, the penetration depth in muscle tissue could be estimated to at least 1 cm (all information provided by the instrument manufacturer). The power level at sample surface was measured to be approx. 210 mW with LaserCheck (Coherent Inc., USA). Each sample was moved under the probe to scan most of the sample surface, as shown in Fig. 1. Pork samples was
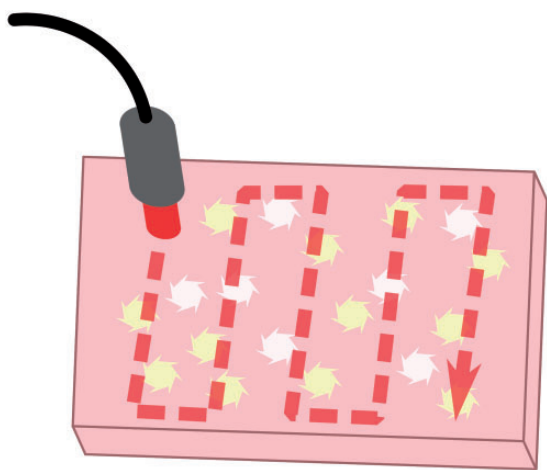
Figure 1. Illustration of scanning path with Raman probe.

scanned in the spectral range from 600 to 1800 cm^{-1} with a resolution of 1 cm^{-1} and an acquisition time of 60 s (four acquisitions of 15 s), where one spectrum was recorded from each sample. Poultry by-products were scanned in the range from 600 to 1800 cm^{-1} with a resolution of 0.3 cm^{-1} and an acquisition time of 80 s (four acquisitions of 20 s). Salmon by-products were scanned in the range from 600 to 1800 cm^{-1} with a resolution of 1 cm^{-1} and an acquisition time of 60 s (12 acquisitions of 5 s). Three spectra were recorded from each poultry and salmon by-product sample. Instrument setup and experiment was controlled using iC Raman version X software (Mettler Toledo, Switzerland).

Near-Infrared Spectroscopy

Near-infrared spectra from pork were recorded with a prototype instrument, designed to measure in interactance mode as described by Wold et al.[28] A halogen light source of 50 W was used to illuminate the sample in two rectangular regions of approximately 2 mm × 20 mm. Distance between the two illuminated regions was ~24 mm. The system collected the light that was transmitted into the meat and came up again in a small area of 4 mm × 4 mm between the two illuminated rectangles. The collected light had then traveled to a depth of ~13 mm.[29] The spectra consisted of 20 evenly spaced wavelengths in the region 760–1100 nm. A total of 100 spectra were collected over 2 s. There was no contact between sample and instrument, working distance to sample was 20 cm.

In the case of the poultry and salmon by-product samples, a Perten DA7440 Process NIR Sensor (Perten Instruments, a Perkin Elmer Company, USA) was used to obtain spectra in reflectance mode at a 25 cm working distance. The spectral range was 950–1650 nm with a resolution of 5 nm. The samples were spread out on a board (40 cm × 40 cm) and each spectrum was acquired at an average of 10 s of acquisition while the samples were moved manually under the spectrometer to scan most of the sample surface. Three and four spectra, for poultry and salmon, respectively, were obtained for each sample and a different surface was scanned each time to obtain representative sample spectra. The average spectrum was used for further analysis. The spectra were transformed from reflectance to absorbance units.

Reference Analysis

Homogenized meat and salmon by-product samples were thawed at 4 °C before determination of IMF or total fat content. This was done using a low field nuclear magnetic resonance instrument (@nTEK, Norway) and a one-shot method for estimation of fat content as described by Sorland et al.[30] Approximately 3 g of homogenized sample was transferred to a Teflon nuclear magnetic resonance

(NMR) tube and temperature was adjusted to 40 °C before NMR measurement. From each homogenate, two sub-samples were analyzed with two replicates from each sub-sample. The average value from homogenates was used as IMF or total fat content reference.

The reference measurements for poultry by-products were performed at an external laboratory (ALS Laboratory). Two parallels from each sample were analyzed. The Dumas method,[31] as used for total N and protein content, was determined as 6.25 times the N-total. Fat content was determined with an internal method at the ALS Laboratory based on pulsed NMR. Samples were dried in an oven to determine the moisture content. After that, samples were stabilized at 50 °C and resonance of samples was determined. The fat content was determined automatically by comparing the resonance of the sample with a calibration curve established using the certified olive oil content.

Protein content in salmon by-product was analyzed at a second external laboratory (Eurofins Agro Testing Norway AS, Norway). Nitrogen was determined by a modified Kjeldahl method following Nordic Committee on Food Analysis protocols.[32] A conversion factor of 6.25 was used to calculate protein content from N-total. Fat and protein content were calculated as percentage of total weight for all samples.

Pre-Processing of Spectra and Data Analysis

Pre-processing of spectral data was done to give comparable spectra for further analysis, by reducing or removing the effect of baseline offset, scattering, and particle size effects. Raman spectra were baseline corrected and fluorescence background was removed using a modified polyfit method as described by Lieber and Mahadevan-Jansen,[33] which is an iterative method for fitting a polynomial curve as baseline. To compare results from baseline correction and normalization, Raman spectra were subjected to standard normal variate (SNV) transformation after baseline correction.[34] NIR spectra from each sample were averaged before pre-processing by applying SNV along the entire spectrum.

Partial least squares regression (PLSR) was used for determining relationship between fat or protein content and spectroscopic data. PLSR extracts information in the spectra that is important for explaining variation in the reference measurements when making models.[35] PLSR models used the entire recorded spectrum for RS and NIRS to make models and were validated using leave-one-out cross-validation.

The ratio of prediction to deviation (RPD) value gives a quick appraisal of a model and is calculated as the standard deviation (SD) of the reference values divided by the models RMSECV.[36] The RPD statistic is a useful tool to compare model performance when reference measurements vary for different models, which makes values for R^2 and RMSECV difficult to compare directly. An RPD value larger than 2.0 is recommended for rough screening purposes, values above 3.0 are considered good to very good, while values over 4.1 are considered excellent and can be used for any application. RPD values below 2.0 are not recommended for prediction purposes. Coefficient of variation (CV) was calculated as SD divided by the mean for a given variable, which shows variation in relation to the mean in the data set and is an unbiased way to compare variations in different data sets.

Pre-processing of Raman spectra was done in Matlab v.R2016b (The MathWorks Inc.). Pre-processing of NIR spectra and data analysis was carried out using The Unscrambler v.11.0 (CAMO Analytics AS, Norway).

Results and Discussion

Reference Measurements

Total fat and protein content were determined in poultry and salmon samples, while only fat content was analyzed in pork samples (Table I). All sample sets showed good variation in fat content, with CV at 0.20 or higher. Variation in protein content was relatively small for salmon samples, evidenced by a CV of only 0.08, but it was higher for poultry samples. Pearson correlation coefficient between fat and protein content was −0.82 and −0.24 for poultry and salmon, respectively. The high correlation for poultry samples might make it difficult to separate the effect of fat content on the models for protein content and vice versa, while this should be insignificant for the salmon

Table I. Reference statistics of pork and by-products from poultry and salmon. All values are expressed as percentage of total weight, except N and CV, which are unitless.

Sample		N	Mean	Min	Max	Range	SD	CV
Poultry	Fat	52	21.22	10.95	43.25	32.30	8.08	0.38
	Protein	52	15.88	9.49	24.60	15.11	3.20	0.20
Pork (Int and Hom)	Fat	99	4.67	1.44	12.96	11.52	2.17	0.46
Salmon	Fat	55	25.89	18.03	35.17	17.14	5.24	0.20
	Protein	55	15.61	12.40	17.90	5.50	1.27	0.08

N: number of samples; SD: standard deviation in %; CV: coefficient of variation (ratio of SD to the mean); Int: intact; Hom: homogenized.

samples. Taken together, all sample sets seem to be suitable for modeling against spectroscopy, with a possible exception of protein content for salmon.

Ranking of sample sets in accordance with heterogeneity, from least to most heterogeneous, gives the following order, homogenized pork, poultry by-product, intact pork, and salmon by-product. Poultry by-product was finely minced, making it less heterogeneous than the intact pork. The salmon samples contained coarsely ground by-product from entire fish and were very heterogeneous.

Estimation of Fat Content with Spectroscopy

Figure 2 shows baseline corrected Raman spectra colored in accordance with the measured fat content. Visual inspection showed a strong relationship between Raman intensity and fat content for poultry by-product (Fig. 2a) and homogenized pork (Fig. 2b), while the connection was less clear for salmon by-product (Fig. 2c). The most pronounced peaks in the Raman spectra from poultry and pork were the ones assigned to saturated regions of fatty acids, e.g., Raman peak at $1060\,cm^{-1}$, assigned to out-of-phase aliphatic C–C stretch, and peaks at ~1300 and $1440\,cm^{-1}$, assigned to methylene twisting and scissoring deformations (CH_2), respectively.[4] In spectra from salmon, peaks assigned to unsaturated regions of fatty acids increased in intensity relative to spectra from poultry and pork, e.g., peak $1270\,cm^{-1}$, which can be assigned to in-plane olefinic hydrogen bend (=CH).

The performance of models for estimation of fat content in the three sample sets are summarized in Table II.
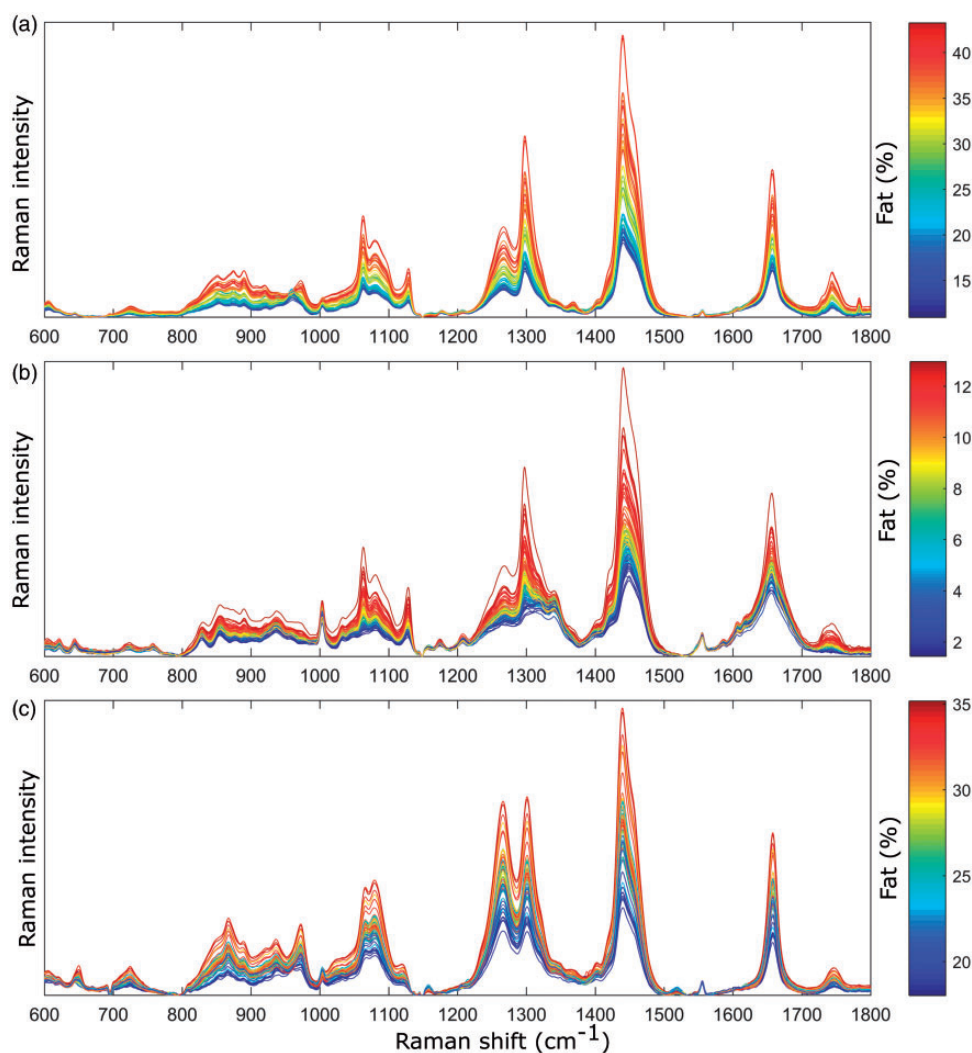
Figure 2. Baseline corrected Raman spectra from (a) poultry by-product, (b) homogenized pork, and (c) salmon by-product used for making the PLSR models. Spectra are colored in accordance with the corresponding fat content of the sample, where fat content in percentage is shown in the color bar for each individual plot.

Table II. Summary of PLSR model statistics for fat and protein content in poultry and salmon by-products and intact and homogenized pork.

	Raman				NIR			
	R^2_{CV}	RMSECV	LV	RPD	R^2_{CV}	RMSECV	LV	RPD
Poultry (fat)	0.94	1.92	2	4.21	0.96	1.72	1	4.70
Pork, Hom (IMF)	0.96	0.43	2	5.05	0.94	0.51	5	4.25
Pork, Int (IMF)	0.73	1.13	2	1.92	0.78	1.01	4	2.15
Salmon (fat)	0.85	2.06	3	2.54	0.89	1.72	2	3.05
Poultry (protein)	0.92	0.94	4	3.40	0.91	0.99	4	3.23
Salmon (protein)	0.56	0.85	3	1.55	0.63	0.79	7	1.61

RMSECV: root mean square error of cross-validation, expressed in %, LV: latent variables in the PLSR model; Hom: homogenized; Int: intact.

Raman models for poultry and homogenized pork were excellent, the model for salmon was good, while the model for intact pork was poor based on their respective RPD values. Results were similar for NIR, but NIR gave slightly better results for all models except for homogenized pork. Even though the model for intact pork is not very good, it shows a substantial improvement compared to previously published studies using RS to estimate IMF in meat.[7,9] The result for intact pork is comparable to those reported in Andersen et al.[10] for RS, but the IMF range was very limited in that study (0.8 to 1.6%), and the current study shows that models hold true over larger spans of IMF. Furthermore, results for homogenized pork show that with representative sampling it was possible to achieve excellent models with RS, encouraging further research to improve sampling techniques.
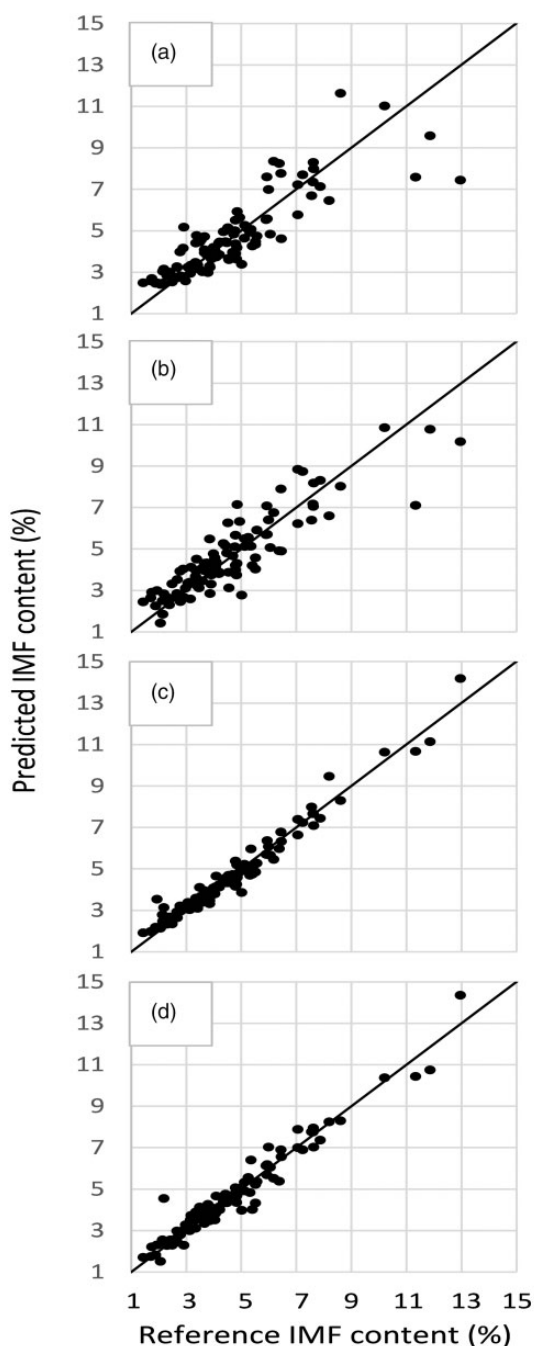
In Fig. 3, predicted values from PLSR models and reference analyses are shown for Raman and NIR for intact and homogenized pork. Results for RS and NIRS were similar for both intact and homogenized pork. It was evident that the samples with highest fat content was most deviant for intact pork, which may be caused by non-representative sampling where some of the fat was hidden from the spectroscopic probe.

Regression coefficients for fat content models revealed that models for intact and homogenized pork were very similar and comparable to the regression coefficient for poultry by-product (Fig. 4). These three regression coefficients were dominated by Raman regions assigned to saturated carbons in fatty acids, e.g., Raman regions at ~1300 and 1440 cm^{-1}, assigned to methylene twisting and scissoring deformations (CH_2), respectively.[4] The regression coefficients for salmon by-product were clearly different from the other three, which can be caused by higher amounts of unsaturated fatty acids in the salmon samples. For instance, the positive peak in the salmon regression coefficient at approx. 1270 cm^{-1} can be assigned to in-plane olefinic hydrogen bend (=CH), and there was a more pronounced peak at approx. 1660 cm^{-1}, which can be assigned to olefinic stretch (C=C).[4]

Estimation of Protein Content Using Spectroscopy

A summary of PLSR model performance for estimation of protein content in poultry and salmon by-products is presented in Table II. Models were very good for poultry by-product but poor for salmon by-product. RS and NIRS gave similar results for both poultry and salmon, with Raman offering marginally better model for poultry, while NIR was slightly better for salmon. In general, the PLSR models for protein content were worse than those for fat content, when judged by their RPD values. It is well known that muscle proteins provide lower Raman signals than lipids. The Raman scattering cross section of the CH_2 bend (around 1440 cm^{-1}) has been estimated, showing up to sixfold increase going from pure proteins to lipid components.[37] In addition, the optical scattering coefficient is higher for adipose tissue than for lean muscle tissue.[38] Both features could most likely explain why protein content models were worse than corresponding fat content models. Another reason could be the lower CV for protein content, especially for salmon.

Regression coefficients for the protein model from poultry were almost a mirror image of the regression coefficients from the fat content model (Fig. 5), probably a result of the relatively high correlation between fat and protein content. Negative peaks were mostly associated with fatty acids, while many of the positive peaks probably can be assigned to protein features, e.g., the broad amide I and III bands at approx. 1645 to 1670 cm^{-1} and 1220 to 1300 cm^{-1}, respectively.[39] Regression coefficients were visibly different between poultry and salmon, where the most apparent difference was the peak at approx. 1660 cm^{-1}, having opposite signs. This might be a consequence of the overlap in the Raman spectrum for amide I and fatty acid features, giving different outcomes for poultry and salmon. Another important protein feature in the salmon regression coefficient plot was that the amino acid phenylalanine, at ~1004 cm^{-1}, also referred to as a probe for protein content.[40] Two other peaks in the salmon regression

Figure 3. Predicted and reference values from cross-validated PLSR models for IMF content in pork. (a) RS intact pork, (b) NIRS intact pork, (c) RS homogenized pork, and (d) NIRS homogenized pork.

coefficient plot stand out as different from the poultry one, namely peaks at approx. 1160 cm^{-1} and 1520 cm^{-1}, assigned to C–C single and double bonds, respectively, in carotenoids.[41] Carotenoids can be found in salmon muscle, which is rich in protein, and carotenoids are therefore likely to be correlated with total protein content.

## Effect of Pre-Processing on Regression Performance

All regression models in the previous sections were obtained on baseline corrected Raman spectra as displayed in Fig. 2. Several studies show the importance of a normalization step in the pre-processing of Raman spectra, to account for, e.g., differences in sample presentation, focusing or sampling volumes, instrument drift (i.e., laser intensity fluctuations), and to aid for calibration transfer between instruments when conducting analysis outside of a laboratory setting over time.[42,43] In many cases, normalization will result in a simplification of the regression models (i.e., fewer PLS components used).[44] Table III compares regression results for baseline corrected spectra with baseline corrected and normalized (SNV) spectra. For the fat content models, models pre-processed with SNV performed equally or poorer compared to only background correction. Also, contrary to the studies mentioned above, there is no clear effect of model simplification by performing normalization. Clearly, as seen in Fig. 2, there is a high correlation between the overall intensity levels and the fat contents. When normalization is performed on such spectra, one obviously stands the risk of removing some of this information. The result of this is shown in Fig. 6 for IMF of homogenized pork, where the spectra that have undergone SNV transformation show a distinct trend in underestimating low and high IMF values. This effect will most likely depend on the span of lipid contents in the sample set. For proteins, it is not easy to see a clear difference between normalized and baseline corrected spectra from Table III. Thus, more data collected over time is needed to draw definitive conclusions about the need for normalization when quantifying bulk properties from Raman spectra.

General discussion

Even though RS performed comparable to NIR in this study, there are still needs to improve sampling and instrumental shortcomings of RS. Firstly, even with wide area Raman probes, the spatial area for recording spectra is still limited. This may not be a problem for semi-heterogeneous discrete units, as representative average spectra can be collected over a large sample area using a scanning approach as in the current study. Secondly, collection of high-quality Raman spectra can be time consuming, independent of sample size. In the current study, standard exposure times based on previous experience was applied, resulting in rather long total exposure times (i.e., from 60 to 80 s) compared to acquisition times of NIR spectra (from 2 to 10 s). Measurement speed can be a limitation for applications where separate samples need to be analyzed real time, for instance single samples passing by on a conveyor belt. When analyzing continuous sample flows where the overall composition over large sample volumes is of
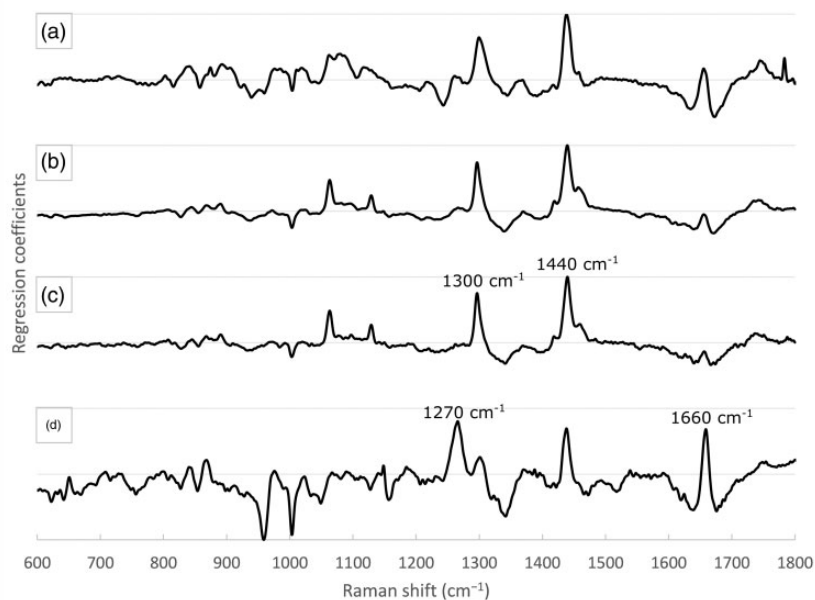
Figure 4. Regression coefficients for total fat models from RS. (a) Poultry by-product, (b) homogenized pork, (c) intact pork, and (d) salmon by-product.
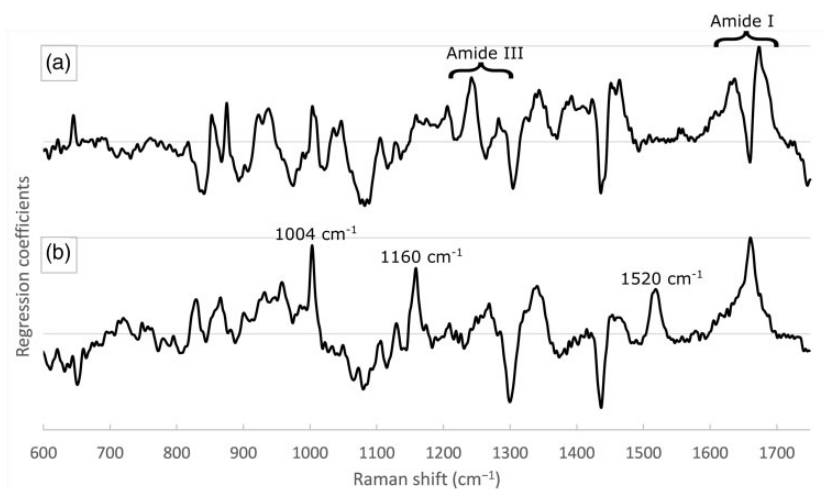

Figure 5. Regression coefficients for protein models from RS. (a) Poultry and (b) salmon by-products.

Table III. PLSR models for fat and protein content from Raman spectra with different pre-processing procedures, only background correction and background correction followed by SNV.

	BC			BC + SNV		
	R^2_{CV}	RMSECV	LV	R^2_{CV}	RMSECV	LV
Poultry (fat)	0.94	1.92	2	0.92	2.28	2
Pork, Hom (IMF)	0.96	0.43	2	0.90	0.69	1
Pork, Int (IMF)	0.73	1.13	2	0.74	1.13	1
Salmon (fat)	0.85	2.06	3	0.83	2.16	3
Poultry (protein)	0.92	0.94	4	0.91	0.98	5
Salmon (protein)	0.56	0.85	3	0.59	0.81	3

RMSECV: root mean square error of cross-validation, expressed in %; LV: latent variables in the PLSR model; BC: baseline corrected; Hom: homogenized; Int: intact.
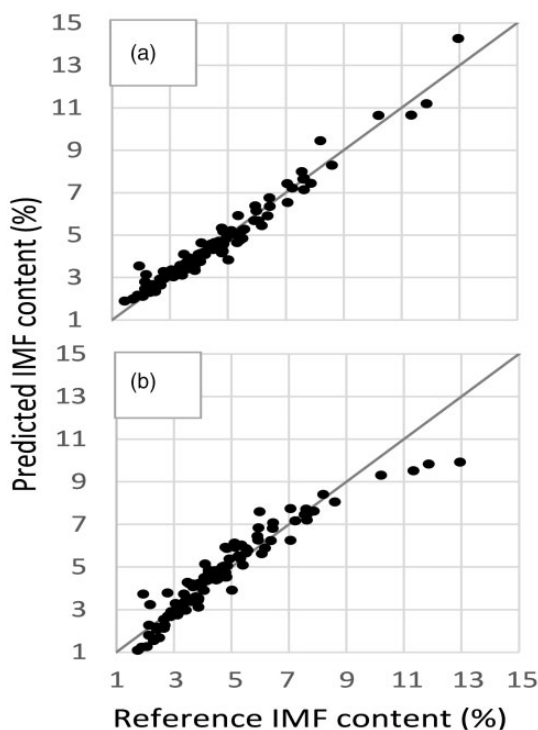
Figure 6. Predicted and reference values for (a) baseline corrected Raman spectra and (b) baseline and SNV corrected Raman spectra for PLS models of IMF content of homogenized pork.

interest, on the other hand, measurement speed is not as important. Optimization in exposure times for collection of adequate quality Raman spectra in industrial settings will be needed and requires attention in future research. This may be helped along by contemporary improvement in Raman instrument design.

As pointed out by Wikström et al.,[22] there are two main benefits with the wide area Raman probe used in the current study: (i) enhanced sampling volume and (ii) robustness towards fluctuations in sampling distances. The former point has already been discussed, but the significance of the latter point should not be forgot when it comes to practical use of RS for food quality applications. In industrial environments, there is simply no time for fine adjustments of the focusing distance. In addition, since we have a system with almost constant sampling volume, this might explain why normalization apparently is not needed for extracting quantitative data on bulk chemical composition in the current study. It could be hypothesized, however, that if minor chemical components, i.e., fatty acid contents, are to be modeled from such spectra, a normalization procedure would be essential to extract this information in a quantitative way, due to the large intensity variations in the bulk chemistry bands. Taking these findings into account, we therefore are starting to see the contours of a Raman system for multicomponent analysis of heterogeneous food samples. This could pave the way for future commercial solutions for RS in the food industry, particularly in

applications where current solutions based on for instance NIRS have limitations in chemical resolution (e.g., assessment of fatty acid composition).

Conclusion

In the present study, RS was shown to successfully estimate selected bulk properties of different food and food by-product matrices utilizing a wide area Raman probe in conjunction with surface scanning, thus giving a representative spectrum for each sample. This shows that RS can fulfill the role as a tool for routine and continuous analysis of foods in industrial relevant applications. Optimization of the technique, both related to instrumental and practical implementation, is still needed to unlock the full potential of RS in food analysis.

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.

ORCID iD

Petter Vejle Andersen https://orcid.org/0000-0002-8241-1051

References

1. Y. Ozaki, R. Cho, K. Ikegaya, et al. "Potential of Near-Infrared Fourier-Transform Raman Spectroscopy in Food Analysis". Appl. Spectrosc. 1992. 46(10): 1503–1507.
2. E.C.Y. Li-Chan. "The Applications of Raman Spectroscopy in Food Science". Trends Food Sci. Technol. 1996. 7(11): 361–370.
3. K.A. Esmonde-White, M. Cuellar, C. Uerpmann, et al. "Raman Spectroscopy as a Process Analytical Technology for Pharmaceutical Manufacturing and Bioprocessing". Anal. Bioanal. Chem. 2017. 409(3): 637–649.
4. J.R. Beattie, S.E.J. Bell, C. Borgaard, et al. "Prediction of Adipose Tissue Composition Using Raman Spectroscopy: Average Properties and Individual Fatty Acids". Lipids. 2006. 41(3): 287–294.
5. O. Monago-Maraña, J.P. Wold, R. Rødbotten, et al. "Raman, Near-Infrared, and Fluorescence Spectroscopy for Determination of

Collagen Content in Ground Meat and Poultry By-Products". LWT Food Sci. Technol. 2021. 140: 110592.

6. J.I. Hikima, M. Ando, H.O. Hamaguchi, et al. "On-Site Direct Detection of Astaxanthin from Salmon Fillet Using Raman Spectroscopy". Mar. Biotechnol. 2017. 19(2): 157–163.

7. S.M. Fowler, E.N. Ponnampalam, H. Schmidt, et al. "Prediction of Intramuscular Fat Content and Major Fatty Acid Groups of Lamb *M. Longissimus Lumborum* Using Raman Spectroscopy". Meat Sci. 2015. 110: 70–75.

8. Y.Q. Nian, M. Zhao, C.P. O'Donnell, et al. "Assessment of Physico-Chemical Traits Related to Eating Quality of Young Dairy Bull Beef at Different Ageing Times Using Raman Spectroscopy and Chemometrics". Food Res. Int. 2017. 99: 778–789.

9. R. Cama-Moncunill, J. Cafferky, C. Augier, et al. "Prediction of Warner–Bratzler Shear Force, Intramuscular Fat, Drip-Loss, and Cook-Loss in Beef via Raman Spectroscopy and Chemometrics". Meat Sci. 2020. 167: 108157.

10. P.V. Andersen, J.P. Wold, E. Gjerlaug-Enger, et al. "Predicting Post-Mortem Meat Quality in Porcine Longissimus Lumborum Using Raman, Near Infrared and Fluorescence Spectroscopy". Meat Sci. 2018. 145: 94–100.

11. K. Shin, H. Chung. "Wide Area Coverage Raman Spectroscopy for Reliable Quantitative Analysis and Its Applications". Analyst. 2013. 138(12): 3335–3346.

12. P. Matousek, A.W. Parker. "Bulk Raman Analysis of Pharmaceutical Tablets". Appl. Spectrosc. 2006. 60(12): 1353–1357.

13. P. Matousek, I.P. Clark, E.R. Draper, et al. "Subsurface Probing in Diffusely Scattering Media Using Spatially Offset Raman Spectroscopy". Appl. Spectrosc. 2005. 59(4): 393–400.

14. N.K. Afseth, M. Bloomfield, J.P. Wold, et al. "A Novel Approach for Subsurface Through-Skin Analysis of Salmon Using Spatially Offset Raman Spectroscopy (SORS)". Appl. Spectrosc. 2014. 68(2): 255–262.

15. J. Qin, M.S. Kim, K. Chao, et al. "Line-Scan Raman Imaging and Spectroscopy Platform for Surface and Subsurface Evaluation of Food Safety and Quality". J. Food Eng. 2017. 198: 17–27.

16. K. Shin, H. Chung, C.W. Kwak. "Transmission Raman Measurement Directly Through Packed Corn Kernels to Improve Sample Representation and Accuracy of Compositional Analysis". Analyst. 2012. 137(16): 3690–3696.

17. M.V. Schulmerich, M.J. Walsh, M.K. Gelber, et al. "Protein and Oil Composition Predictions of Single Soybeans by Transmission Raman Spectroscopy". J. Agric. Food Chem. 2012. 60(33): 8097–8102.

18. R. Singh, T.P. Wrobel, P. Mukherjee, et al. "Bulk Protein and Oil Prediction in Soybeans Using Transmission Raman Spectroscopy: A Comparison of Approaches to Optimize Accuracy". Appl. Spectrosc. 2019. 73(6): 687–697.

19. R.J. Beattie, S.J. Bell, L.J. Farmer, et al. "Preliminary Investigation of the Application of Raman Spectroscopy to the Prediction of the Sensory Quality of Beef Silverside". Meat Sci. 2004. 66(4): 903–913.

20. J.Y. Ma, D. Ben-Amotz. "Rapid Micro-Raman Imaging Using Fiber-Bundle Image Compression". Appl. Spectrosc. 1997. 51(12): 1845–1848.

21. M.V. Schulmerich, W.F. Finney, R.A. Fredricks, et al. "Subsurface Raman Spectroscopy and Mapping Using a Globally Illuminated Non-Confocal Fiber-Optic Array Probe in the Presence of Raman Photon Migration". Appl. Spectrosc. 2006. 60(2): 109–114.

22. 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.

23. Y. Kim, S. Lee, H. Chung, et al. "Improving Raman Spectroscopic Differentiation of the Geographical Origin of Rice by Simultaneous Illumination Over a Wide Sample Area". J. Raman Spectrosc. 2009. 40(2): 191–196.

24. P.V. Andersen, N.K. Afseth, E. Gjerlaug-Enger, et al. "Prediction of Water Holding Capacity and Ph in Porcine Longissimus Lumborum Using Raman Spectroscopy". Meat Sci. 2020. 172: 108357.

25. O. Monago-Marana, C.E. Eskildsen, N.K. Afseth, et al. "Non-Destructive Raman Spectroscopy as a Tool for Measuring Asta Color Values and Sudan I Content in Paprika Powder". Food Chem. 2019. 274: 187–193.

26. S.G. Wubshet, J.P. Wold, U. Bocker, et al. "Raman Spectroscopy for Quantification of Residual Calcium and Total Ash in Mechanically Deboned Chicken Meat". Food Control. 2019. 95: 267–273.

27. S.G. Wubshet, J.P. Wold, N.K. Afseth, et al. "Feed-Forward Prediction of Product Qualities in Enzymatic Protein Hydrolysis of Poultry By-Products: A Spectroscopic Approach". Food Bioprocess Technol. 2018. 11(11): 2032–2043.

28. J.P. Wold, M. O'Farrell, J. Tschudi, et al. "In-Line and Non-Destructive Monitoring of Core Temperature in Sausages During Industrial Heat Treatment by NIR Interaction Spectroscopy". J. Food Eng. 2020. 277: 109921.

29. J.P. Wold. "On-Line and Non-Destructive Measurement of Core Temperature in Heat Treated Fish Cakes by NIR Hyperspectral Imaging". Innovative Food Sci. Emerging Technol. 2016. 33: 431–437.

30. G.H. Sorland, P.M. Larsen, F. Lundby, et al. "Determination of Total Fat and Moisture Content in Meat Using Low Field NMR". Meat Sci. 2004. 66(3): 543–550.

31. J.B.A. Dumas. Memoire sur quelques points de la théorie atomistique. Paris: Impr. de C. Thuau, 1831.

32. Nordic Committee on Food Analysis (NMKL). "Nitrogen. Determination in Foods and Feeds According to Kjeldahl. No. 6". 2003. https://www.nmkl.org/index.php/en/publications/item/nitrogen-kjeldahl-nmkl-6 [accessed March 9 2021].

33. C.A. Lieber, A. Mahadevan-Jansen. "Automated Method for Subtraction of Fluorescence from Biological Raman Spectra". Appl. Spectrosc. 2003. 57(11): 1363–1367.

34. R.J. Barnes, M.S. Dhanoa, S.J. Lister. "Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra". Appl. Spectrosc. 1989. 43(5): 772–777.

35. H. Martens, M. Martens. "Analysis of Two Data Tables X and Y: Partial Least Squares Regression (PLSR)". In: H. Martens, M. Martens, editors. Multivariate Analysis of Quality: An Introduction. Chichester, UK: John Wiley and Sons Ltd., 2001. Pp. 111–126.

36. P. Williams. "The RPD Statistic: A Tutorial Note". NIR News. 2014. 25(1): 22–26.

37. R. Manoharan, J.J. Baraga, M.S. Feld, et al. "Quantitative Histochemical Analysis of Human Artery Using Raman Spectroscopy". J. Photochem. Photobiol., B. 1992. 16(2): 211–233.

38. S.L. Jacques. "Optical Properties of Biological Tissues: A Review". Phys. Med. Biol. 2013. 58(11): R37–R61.

39. S. Krimm, J. Bandekar. "Vibrational Spectroscopy and Conformation of Peptides, Polypeptides, and Proteins". Adv. Protein Chem. 1986. 38: 181–364.

40. T.W. Barrett, W.L. Peticolas, R.M. Robson. "Laser Raman Light-Scattering Observations of Conformational Changes in Myosin Induced by Inorganic Salts". Biophys. J. 1978. 23(3): 349–358.

41. H. Henmi, M. Hata, M. Takeuchi. "Resonance Raman and Circular Dichroism Studies of Astaxanthin and/or Canthaxanthin in Salmon Muscle". Nippon Suisan Gakkaishi. 1990. 56(11): 1825–1828.

42. N.K. Afseth, V.H. Segtnan, J.P. Wold. "Raman Spectra of Biological Samples: A Study of Preprocessing Methods". Appl. Spectrosc. 2006. 60(12): 1358–1367.

43. J.R. Beattie, J.V. Glenn, M.E. Boulton, et al. "Effect of Signal Intensity Normalization on the Multivariate Analysis of Spectral Data in Complex 'Real-World' Datasets". J. Raman Spectrosc. 2009. 40(4): 429–435.

44. 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.