

RESEARCH

Open Access



RAMAN AND ATR-FTIR SPECTROSCOPY TOWARDS CLASSIFICATION OF WET BLUE BOVINE LEATHER USING RATIOMETRIC AND CHEMOMETRIC ANALYSIS

Megha Mehta^{1*} , Rafea Naffa¹, Catherine Maidment¹, Geoff Holmes¹ and Mark Waterland²

Abstract

There is a substantial loss of value in bovine leather every year due to a leather quality defect known as “looseness”. Data show that 7% of domestic hide production is affected to some degree, with a loss of \$35 m in export returns. This investigation is devoted to gaining a better understanding of tight and loose wet blue leather based on vibrational spectroscopy observations of its structural variations caused by physical and chemical changes that also affect the tensile and tear strength. Several regions from the wet blue leather were selected for analysis. Samples of wet blue bovine leather were collected and studied in the sliced form using Raman spectroscopy (using 532 nm excitation laser) and Attenuated Total Reflectance - Fourier Transform InfraRed (ATR-FTIR) spectroscopy. The purpose of this study was to use ATR-FTIR and Raman spectra to classify distal axilla (DA) and official sampling position (OSP) leather samples and then employ univariate or multivariate analysis or both. For univariate analysis, the 1448 cm^{-1} (CH_2 deformation) band and the 1669 cm^{-1} (Amide I) band were used for evaluating the lipid-to-protein ratio from OSP and DA Raman and IR spectra as indicators of leather quality. Curve-fitting by the sums-of-Gaussians method was used to calculate the peak area ratios of 1448 and 1669 cm^{-1} band. The ratio values obtained for DA and OSP are 0.57 ± 0.099 , 0.73 ± 0.063 for Raman and 0.40 ± 0.06 and 0.50 ± 0.09 for ATR-FTIR. The results provide significant insight into how these regions can be classified. Further, to identify the spectral changes in the secondary structures of collagen, the Amide I region ($1600\text{--}1700\text{ cm}^{-1}$) was investigated and curve-fitted-area ratios were calculated. The $1648:1681\text{ cm}^{-1}$ (non-reducing: reducing collagen types) band area ratios were used for Raman and $1632:1650\text{ cm}^{-1}$ (triple helix: α -like helix collagen) for IR. The ratios show a significant difference between the two classes. To support this qualitative analysis, logistic regression was performed on the univariate data to classify the samples quantitatively into one of the two groups. Accuracy for Raman data was 90% and for ATR-FTIR data 100%. Both Raman and ATR-FTIR complemented each other very well in differentiating the two groups. As a comparison, and to reconfirm the classification, multivariate analysis was performed using Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). The results obtained indicate good classification between the two leather groups based on protein and lipid content. Principal component score 2 (PC2) distinguishes OSP and DA by symmetrically grouping samples at positive and negative extremes. The study demonstrates an excellent model for wider research on vibrational spectroscopy for early and rapid diagnosis of leather quality.

Keywords: Raman spectroscopy, Attenuated Total reflectance - Fourier transform InfraRed spectroscopy, Principal component analysis, Linear discriminant analysis, Wet blue, Distal axilla, Official sampling position

* Correspondence: megha.mehta@lasra.co.nz

¹NZ Leather and Shoe Research Association (LASRA[®]), Palmerston North, New Zealand

Full list of author information is available at the end of the article

1 Introduction

Every year more than a billion animals are slaughtered as part of the animal production industry for meat. In turn, this generates returns of over a billion dollars for the global leather industry, meat processing's most important co-product sector [1, 2]. The production of leather is split into three phases – animal slaughtering, tanning and manufacturing of the finished product for the commercial market. Tanning is one of the most important stages in the leather production. It involves processing the raw skin or hide, to retain its natural properties by stabilising the molecular structure and make it more durable [3]. Previously, natural chemicals like plant tannins, alum, and other minerals were used in the tanning process which had some advantages over current methods using synthetic chemicals, though these take only a fraction of the processing time required for the earlier methods [4]. Wet Blue refers to part-processed chrome-tanned leather in the wet state. During this stage the skin or hide is protected from decomposition through chemical crosslinking that stabilises the collagen network [5]. The blue colour comes from the chromium tanning agent (Chromium (III) oxide), which stays in the leather after tanning.

Looseness is a fault found in leather that affects the quality of the leather. It manifests itself as corrugations on the outer surface of finished leather when bent inward. Whilst processing is known to exaggerate the fault, the root cause of the less densely packed fibres in affected regions is poorly understood, although potential causes may include environment, nutrition, breed and age. Looseness is a major concern to the leather industry in terms of its effect on structure of the leather and appearance of the final leather product [6]. At present looseness can only be accurately identified once the leather is dried, thus tanners can only address it by either discarding the leather or remedial treatment, costing both time and money [7–9]. There is an understanding that looseness is prevalent in specific regions including the shoulder and flanks whereas other regions such as the backbone and official sampling position (OSP) are typically unaffected. This study investigates wet blue leather from two regions – distal axilla (DA), i.e., from the flank side and official sampling position (OSP), i.e., from near the central lower region. The aim is to obtain a better understanding of how tight and loose wet blue leather might be differentiated through measurement, since the OSP and distal axilla regions typically give tighter or looser leathers, respectively. The intention of this study is to develop a model using non-destructive techniques that can identify the looseness fault at an early stage of the leather production. Different strategies or markets for the affected hides can then be identified, to not only save time but also to minimise the damaging costs incurred by identifying looseness at a later stage of processing.

Vibrational spectroscopy techniques such as Raman spectroscopy and Attenuated Total Reflectance - Fourier Transform InfraRed (ATR-FTIR), supported with ratio-metric band intensity analysis and chemometric methods, are used here to identify structural variations which effect the physical properties of leather from the two identified regions – OSP and DA.

Raman spectroscopy measures the inelastic scattering of photons (with visible wavelengths) while they interact with the vibrational motions of molecules to provide useful information about molecular structure via both band position and intensity. Raman can be used for non-invasive probing of chemical and biological samples [10, 11]. Infrared spectroscopy (IR) is based on the absorbance of infrared photons by molecules due to vibrational motion of the molecules present in the matrix. Both non-destructive techniques are fast, require minimal sample preparation, and have high specificity and sensitivity [12]. Raman spectroscopy has the advantage of a very weak water signal so minimal interference from water in biological samples [13], not causing any damage to the sample [14] and allowing in-situ detection using optical fibres or microscopes. Raman is particularly sensitive to structures that are easily polarised, such as aromatic rings and sulphur-containing groups. Water has an absorption that can mask the characteristic amide I band at 1640 cm^{-1} and a very intense, broad absorption around 3300 cm^{-1} , that can obscure absorption by other O-H and N-H vibrations. If water interference can be minimised, then the advantage of IR is its sensitivity to vibrations associated with the amide bonds in proteins. The secondary and tertiary structures of proteins influence the shape of the amide bands and IR spectroscopy provides useful information about protein structure. We have used an ATR-FTIR spectrometer that limits water interference by using a very short effective path length that results from the attenuated total reflection process.

For most studies of spectral diagnosis of biological samples, the mid-IR (MIR) spectrum within $4000\text{--}600\text{ cm}^{-1}$ range seems to be more effective than the near-IR (NIR) range ($14,000\text{--}4000\text{ cm}^{-1}$). Bands within the range of $4000\text{--}1500\text{ cm}^{-1}$ are characterized by various stretching modes of functional groups of molecules. Bands below 1500 cm^{-1} are dominated by deformation, bending and ring vibrations of the molecular “backbone”, and are generally referred to as the fingerprint region of the spectrum. As the vibrational activity between Raman and infrared (IR) spectroscopies is different, some modes in both are active, but others are only Raman or IR active. MIR and Raman spectra both exhibit amide bands that are relevant to the structure of collagen. Thus, IR and Raman spectroscopies provide similar and complementary information of molecular vibrations [15–17].

Sample preparation is relatively simple as compared to other analytical techniques, such as high-performance liquid chromatography (HPLC) and colorimetric methods [18–22]. Finding the variations in the initial leather processing steps reduces the costs of down-stream processing. Therefore, the label-free and non-destructive techniques are highly attractive tool for understanding wet blue leather [13, 16, 17].

Several bone studies have utilised Raman and IR spectroscopy to identify defects [5, 23, 24], the quality of bone affected by bacteria [25–27], or changes in collagen due to cross-links [20–22, 28, 29], but no work has so far been performed on wet blue leather defects. To the best of our knowledge, this work is the first attempt to identify the variation between loose and tight leather regions using these two techniques.

2 Experimental

2.1 Sample preparation

All bovine wet blue samples were prepared by New Zealand Leather and Shoe Research Association (LASRA[®]) using the conventional methods [19]. Samples were collected from the official sampling position (OSP) and distal axilla (DA) of the wet-blue and stored at below 4 °C until analysis.

Wet blue samples were sliced using a Leica CM1850UV Cryostat to 40 μm thickness. Six replicates of each sample were cut and placed on a microscope slide for Raman and ATR-FTIR analysis as described below.

2.2 Data acquisition and spectral processing

Six leather samples labelled as ‘DA’ and displaying signs of looseness and five wet blue bovine leather samples labelled as ‘OSP’ from the tighter regions of the hide were prepared for analysis using the method described above. These samples were then analysed using a home-built Raman microscope utilising a Teledyne-Princeton Instruments (USA), FERGIE spectrometer using a 532 nm excitation laser (with ~ 10 mW laser power) focused onto the sample with a spot size diameter of ~ 1 – 2 μm using $40\times$ magnification and 0.65 NA objective. For both Raman and IR measurements, spectra were collected from 5 different spots on each sample. Raman spectra were acquired with an exposure time of 5 s per frame and 10 frames (each frame was stored separately). Therefore, 50 spectra were obtained from each DA and OSP sample.

A Thermo Scientific[™] iD5 Nicolet[™] iS[™]5 Attenuated Total Reflectance - Fourier Transform InfraRed (ATR-FTIR) spectrometer was used to collect ATR-FTIR spectra from the same wet blue samples. Spectra were recorded by attenuated total reflection (ATR) on a diamond crystal and 16 scans were collected from 5 different spots for each sample.

Figure 1 shows the flowchart for spectral analysis. For analysis by principal components analysis, each spectrum was preprocessed with an algorithm written using the SciKit Learn package [30] in Python 3.7. Baseline correction, background subtraction and average spectra were obtained using the Python algorithm.

For ratiometric analysis, Origin 2018b (Origin Lab Corporation, Northampton, MA, USA) was used. Pre-processing, consisting of a 7-point, zero-order derivative Savitzky-Golay smoothing function, was applied to smooth spectral noise. Curve fitting by sums-of-Gaussians was used to determine band areas, which were subsequently used to calculate area ratios of the peaks of interest.

3 Results and discussion

Raman and ATR-FTIR spectra are shown in Figs. 2 and 3 respectively. Bands that are known to be associated with functional groups and structures in protein are labelled in the Raman spectra. Bands with positions within instrumental resolution in the Raman and FTIR spectra are assumed to have the same chemical and structural origin [31].

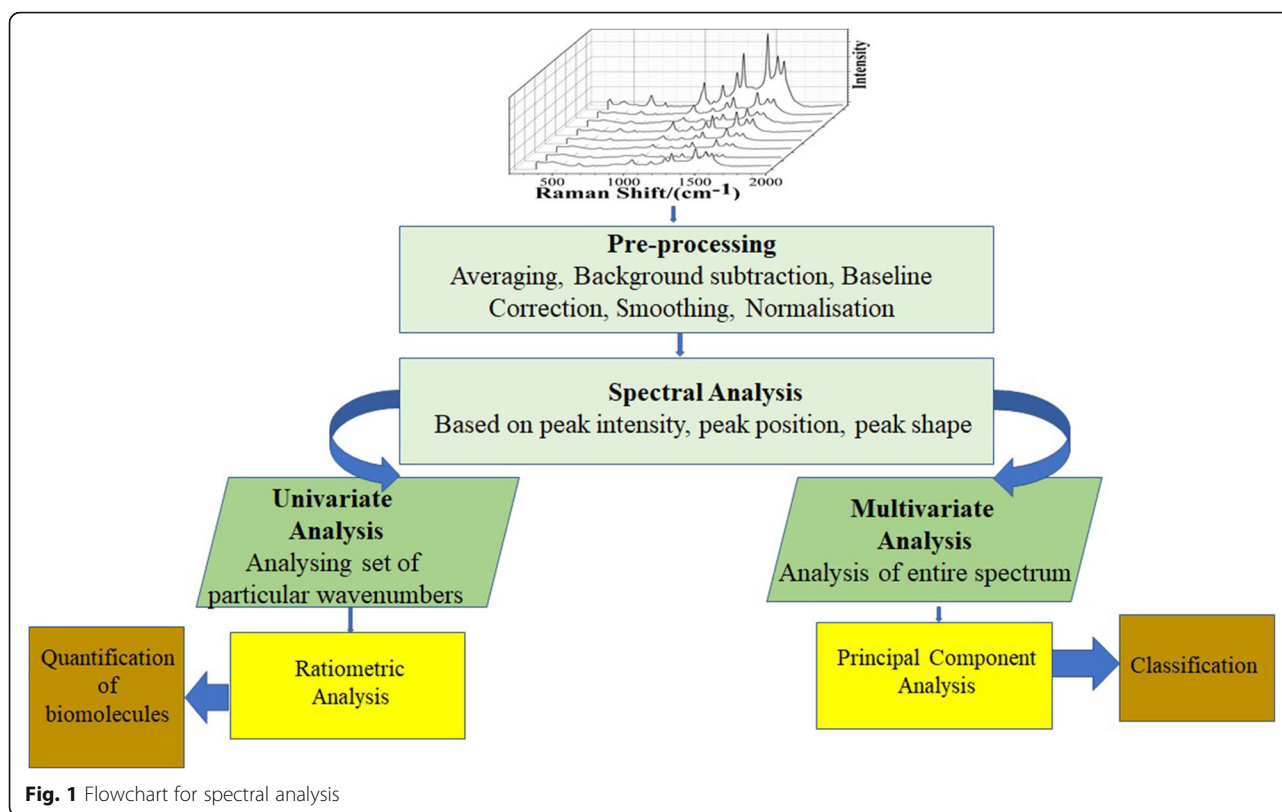
3.1 Peaks of interest

Band assignments and their interpretation are based on Raman and IR studies of collagen tissues.^{16,17,32} By visual inspection, it was found that there are variations in the collagen region (1002 – 1680 cm^{-1}). A careful examination of the spectra showed shifting of a few peaks due to the complexity of biochemical components in leather samples. Sharp peaks were observed in OSP leather Raman spectra whereas significant overlapping of bands was found in DA in the 1550 – 1700 cm^{-1} region.

The observed peak positions of the Raman and IR bands observed and their assignment for wet blue leather are shown in Table 1.

There is a significant shift observed in peak position, intensity and number of signature peaks between DA and OSP samples in Raman and IR spectra. Both DA4 and DA5 show a broad band different from the other DA replicates, which indicates that some structural changes in collagen may occur due to alterations in secondary structures – α helix, β sheet, random coils or immature cross links.

The peak identified at 1669 cm^{-1} is associated with random or unordered protein structure (e.g., random coils). The amide I vibration is dominated by peptide carbonyl stretching vibration with some contribution of C–N stretching and N–H in-plane bending [32]. The bands near 875 and 920 cm^{-1} can be assigned to the C–C stretching vibrations of amino acids characteristic of collagen; hydroxyproline and proline. The band near 1002 cm^{-1} is assigned to the phenyl ring breathing mode of amino acid, phenylalanine [15, 28, 33]. 2340 cm^{-1}



band is observed in few DA and OSP samples is the appearance of asymmetric stretching of CO_2 band which is the result of some background from the spectra.

Special emphasis was placed on the spectral features at 1448 cm^{-1} which is assigned to the CH_2 bend of phospholipids [34] and $1650\text{--}1669\text{ cm}^{-1}$, that corresponds to amide I region that is comprised of both proteins and lipids [35, 36]. Selecting these two bands serves as an excellent indicator of variations, because any changes due to lipid variation are factored out using the 1448 cm^{-1} lipid band [37, 38]. It was found that 1448 cm^{-1} is the more intense Raman band when compared to 1669 cm^{-1} whereas 1632 cm^{-1} is the most intense IR band. Therefore, collagen analysis was performed using the peak area ratio of the CH_2 wag band at 1448 cm^{-1} and the amide I band at 1669 cm^{-1} for Raman analysis.

Before analysing the Raman and IR marker bands for ratiometric analysis, we decided to validate the accuracy of the method and spectral positions identified for DA and OSP. Raman analysis was performed on the other regions of the wet blue to classify loose and tight features based on specific biomarker Raman bands.

Our hypothesis is that OSP region tend to give tight leather as it comes from central backbone part of hide whereas DA region is more prone to looseness as it

comes from flanks or sides of hide. To confirm that the features we have identified are characteristic of looseness and not simply associated with the location of the sample, we have a selected few regions from OSP looking at some wrinkles which exhibit characteristic looseness and a few regions from DA which are not wrinkled or stretched to investigate for tightness.

We have also selected regions from other parts of wet blue like neck and shoulder. The selection of samples was based on the visual examination of wet blue. Figure 4 shows the Raman spectra of all these regions labelled as OSP (tight); the characteristic feature of OSP, OSP (loose) means few loose regions in OSP, DA (loose); the characteristic feature of DA and DA (tight) meaning any tight regions observed in DA.

Table 2 summarises some differentiating Raman bands from OSP, DA and other regions of wet blue leather. A few characteristic bands categorise wet blue into two sets as loose and tight leather rather than OSP and DA. The following two signals are of interest in classifying looseness:

- (1) the protein backbone confirmation: the amide I band detected at 1677 cm^{-1} for tight regions in OSP and DA corresponds to anti-parallel β -sheets which give a tight structure whereas the amide I

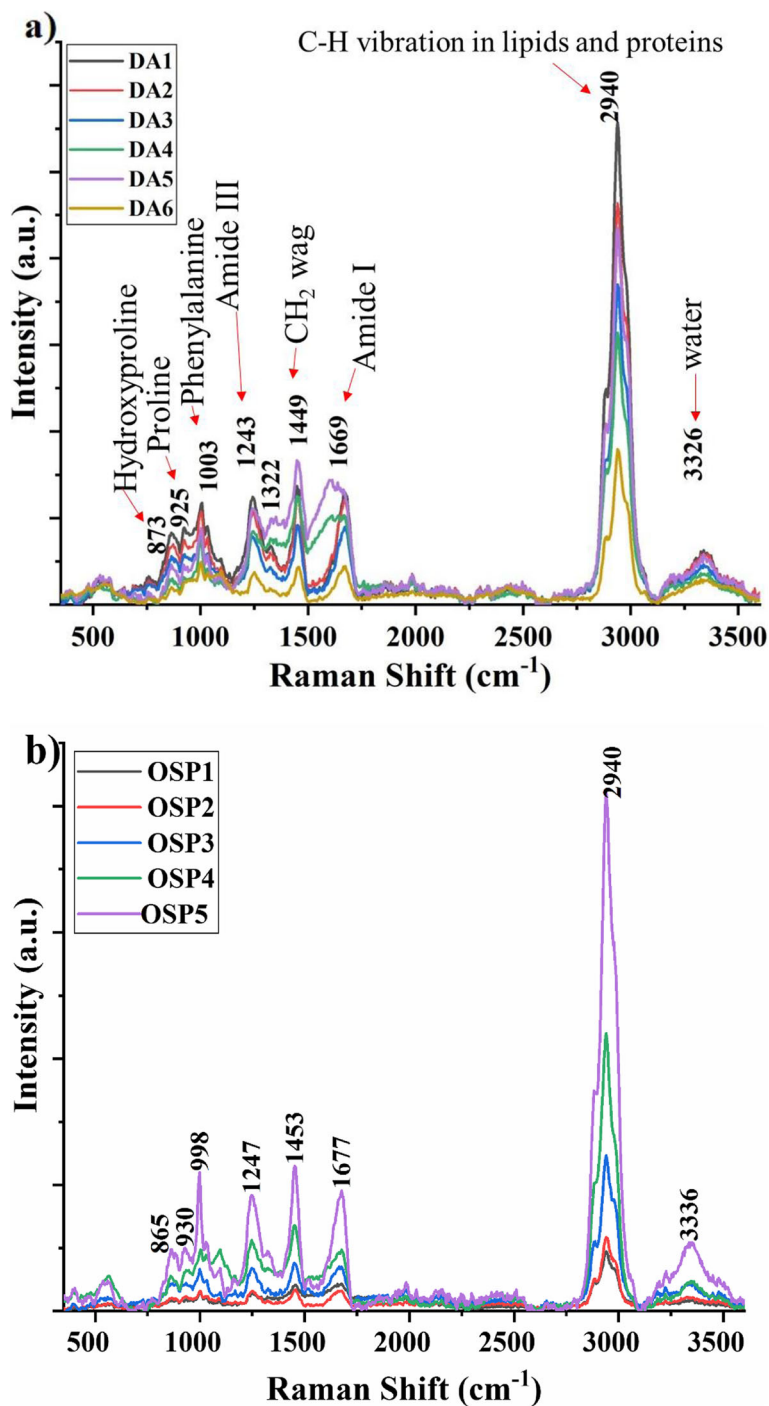


Fig. 2 Raman spectra of (a) loose and (b) tight leather samples

band detected at 1665 or 1669 cm^{-1} belongs to disordered structures associated with looseness.
 (2) Another interesting common signal found in the loose regions of OSP and DA is the Raman peak at 1326 cm^{-1} which is attributed to CH_3CH_2 wagging mode of collagen which is more enhanced in loose regions of DA. The phenylalanine Raman band at

1002 cm^{-1} is common in loose regions of DA and OSP. Raman bands from other regions of wet blue have similar characteristic peaks as OSP tight which differentiates it from loose parts of wet blue. Because the 1677 cm^{-1} is a marker band for tight and 1665–1669 cm^{-1} is a marker band for loose, we can find that OSP does have loose regions and DA

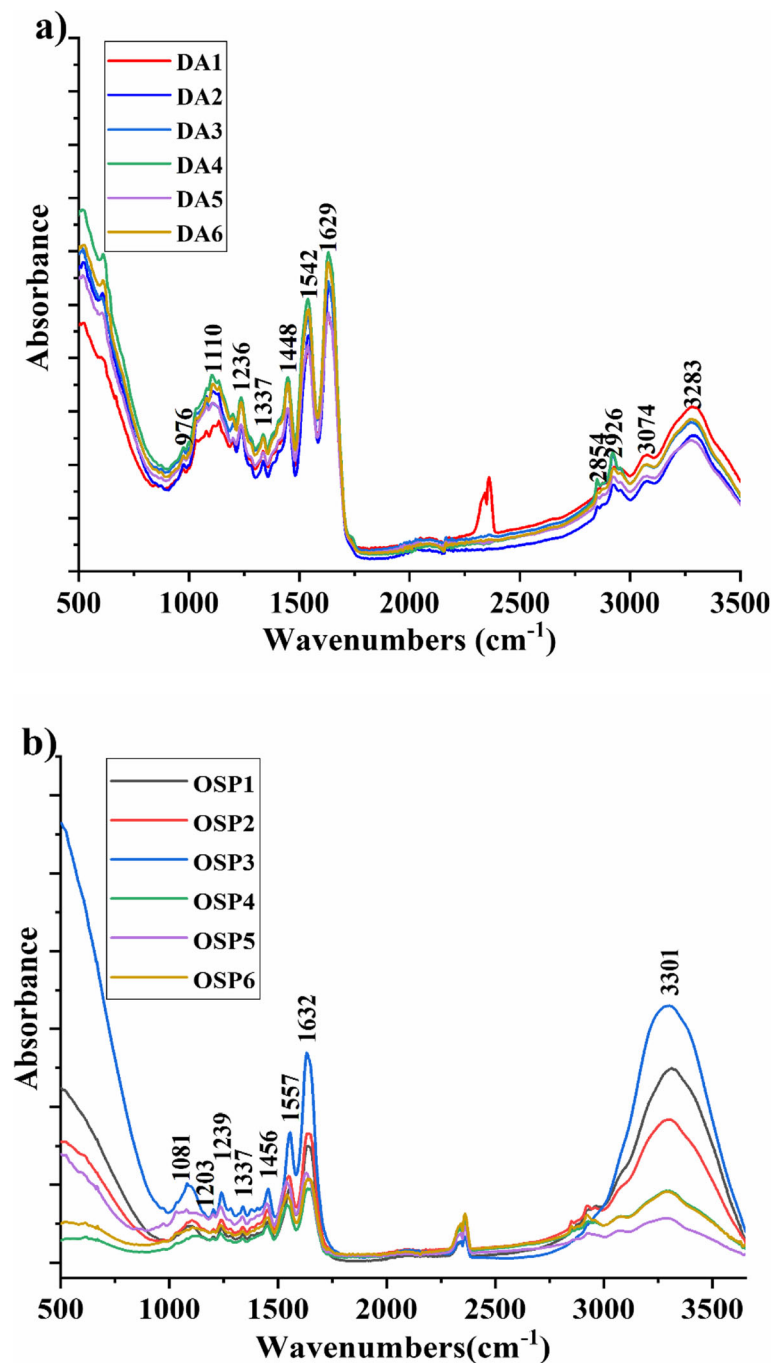


Fig. 3 ATR-FTIR spectra of (a) loose and (b) tight leather samples

has tight regions. These marker bands are indeed classifying the leather based on their structure (tight vs loose) and not merely on location (DA vs OSP).

But the loose and tight features are more dominant in DA and OSP regions and visible in the spectra (Figs. 2,3, and 4), despite of looking specifically loose regions in

tight section of OSP or tight regions in loose section of DA, therefore has the potential of further classification based on location. Hence, further study was carried out in analysing DA and OSP locations for classifying loose and tight structural features to demonstrate the potential of Raman and ATR-FTIR spectroscopy.

Two spectral analysis techniques were employed to find the best classification fit for various biochemical

Table 1 Raman and infrared spectroscopic band assignments for Wet Blue leather

Peak position (cm ⁻¹)	Assignment	Comments	Observable
1002	v (C-C)	Phenylalanine referred to Collagen	Raman
1247	Amide III	Protein β sheet and α helix material	IR, Raman
1448/1449	γ (CH ₂) bending	CH ₂ deformation of protein and lipids	IR, Raman
1550–1596	Amide II	N-H bend + C-H stretch (Proteins)	IR
1629–1677	Amide I Complex	C=O stretch also relates to Collagen, Elastin, β sheet and immature cross-links.	IR, Raman
3200–3300	O-H stretching vibrations	water content	IR, Raman

components affecting the wet blue leather quality, such as proteins, lipids or nucleic acids.

- A univariate statistical method which includes band intensities, area ratios and intensity ratio calculations for the interpretation of spectra [42, 43]. This ratiometric analysis was carried out for qualitative classification, which was further supported by a logistic regression algorithm to enable straightforward quantitative classification of DA and OSP.
- A multivariate statistical method (based on Principal Component Analysis) that considers the whole spectrum but performs classification with a small number of variables (data set dimension reduction) that extract the maximum variance in the data. Multivariate analysis makes no a priori assumptions about selecting the best variables for classification.

3.2 Univariate analysis

Ratiometric analysis, a simple approach, was employed to identify the spectral variations by Raman and IR spectroscopy and generate a systematic and comparative trend of structural features of biochemical components in OSP and DA. Ratiometric analysis can overcome variations due to sample thickness and morphology, background scattering fluctuations and other instrumental effects [17].

Intensity-based ratiometric analysis may result in inaccurate interpretation due to baseline estimation issues [35, 44]. So, average and standard deviations of the peak area ratios for the CH₂ deformation (1448 cm⁻¹) and Amide I (1669 cm⁻¹) bands from Figs. 2 and 3 were calculated [16, 17]. For the Raman spectra, 0.57 ± 0.099 and 0.73 ± 0.063 for DA and OSP ratio values were obtained respectively. For IR, the values were 0.40 ± 0.057 and 0.49 ± 0.13 for DA and OSP samples (Additional file 1). Both Raman and IR show significant variation between the two categories of loose and tight samples. Although ATR-FTIR and Raman spectroscopy arise from the same physical phenomenon of molecular vibration, the processes of Raman scattering, and infrared absorption are fundamentally different as observed in Table 1. Additional bands like 1548 cm⁻¹, observed in IR, but absent

in Raman provides an understanding of cross-links in collagen [44, 45]. Hence, the combination of Raman and FT-IR gives synergistic information on complex samples in a non-destructive manner.

The variation between the two categories could be the result of changes in the collagen network, which directs further investigation towards the amide I band of collagen, which consists of several secondary structures [24]. *Curve fitting by sums-of-Gaussians* method was used to find the component area under a broad band. Accurate peak areas, and peak centres then can be deduced. Univariate analysis was again performed using the collagen components in the section below.

3.2.1 Alterations in collagen network

The amide regions of proteins are overlapped by many underlying bands [27]. In vibrational spectroscopic methods, such as FTIR and Raman, resolution of underlying constituent peaks and calculation of their contributions offer a wealth of information, as these peaks are very sensitive to secondary structure [15, 46].

Therefore, curve-fitting was carried out on both Raman and IR data to investigate the spectral changes in the secondary structures of collagen. A typical result of curve fitting four Gaussian components to the Amide I band in the Raman spectrum is shown in Fig. 5. These secondary bands have been used to investigate the lipid to protein ratio as a measure of collagen quality.

There is a well-established frequency – assignment correlation in literature (Table 3) for the underlying bands in amide I group [34, 43]. The Amide I band in the OSP spectrum is strongly asymmetric and its curve-fitting (Fig. 5) yields components in the 1600–1700 cm⁻¹ region which can be mainly assigned to collagen (1648 and 1669 cm⁻¹), elastin (1681 cm⁻¹), and amino acids (1610 and 1698 cm⁻¹).

The triple helical structure of the collagen molecule is unique and there is no specific peak wavenumber for these secondary structures (e.g., α-helix, or β-sheet). Therefore, changes in collagen's helical structure were investigated empirically by observing changes in the curve fitted area ratios, as identified by the curve-fitting analysis.

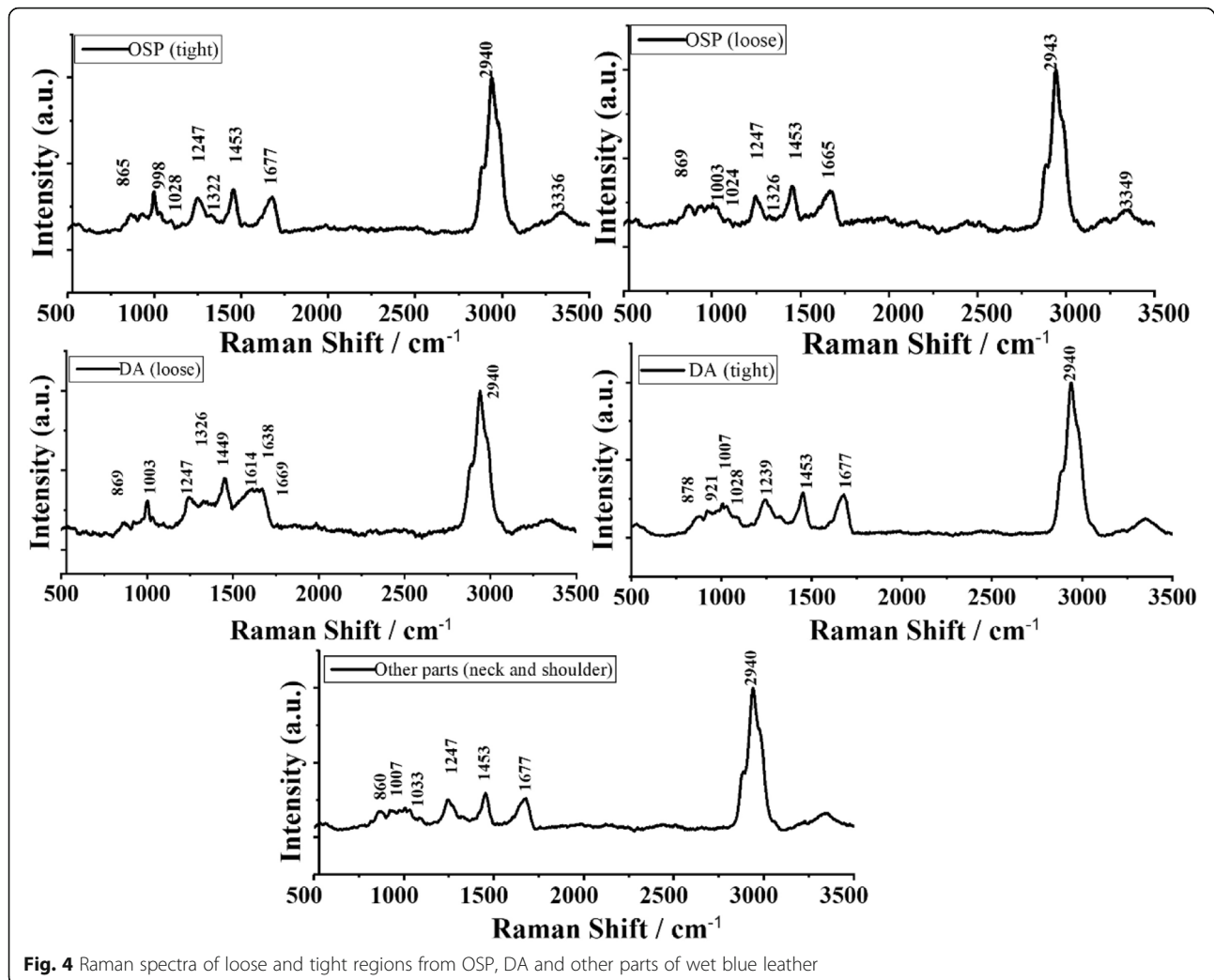


Fig. 4 Raman spectra of loose and tight regions from OSP, DA and other parts of wet blue leather

Table 2 Major Raman band assignments distinguishing loose and tight wet blue leather [16, 17, 39–41]

Peak position (cm ⁻¹)	Assignments	Chemical groups	Observable
1003	v (C-C) skeletal; phenylalanine	Amino acids	DA (loose), OSP (loose)
1028	v (C-C) skeletal; phenylalanine of collagen	Amino acids	OSP (tight), DA (tight)
1247	Amide III	Proteins	DA (loose), OSP (loose), OSP (tight)
1239	Amide III	Proteins	DA (tight)
1326	CH ₃ CH ₂ wagging mode present in collagen.	Amino acids	DA (loose), OSP (loose)
1449	C-H vibration	Proteins and Lipids	DA (loose)
1453	asymmetric CH ₃ bending & CH ₂ scissoring (is associated with elastin, collagen, & phospholipids)	Proteins and Lipids	OSP (loose), OSP (tight), DA (tight)
1677	Anti-parallel β-sheet; Amide I	Proteins	OSP (tight), DA (tight)
1665–1669	Disordered structure; Amide I	Proteins	DA (loose), OSP (loose)

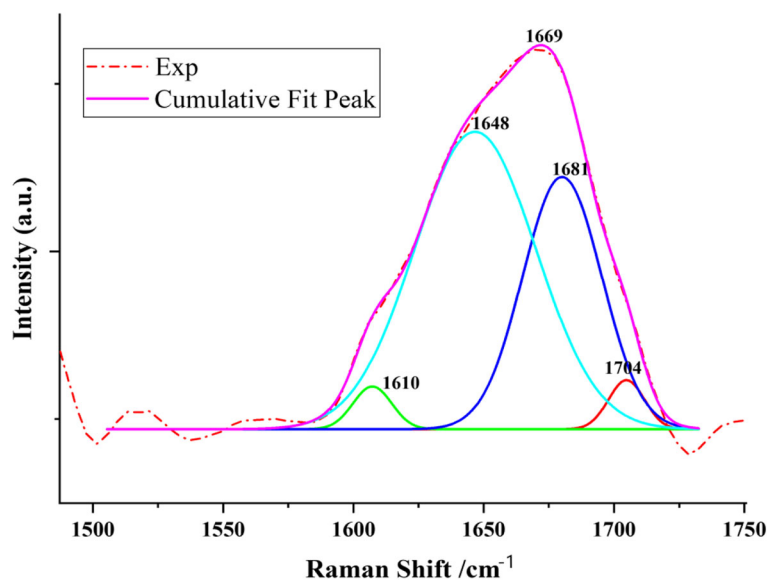


Fig. 5 Typical curve fitting of Amide I Raman band centred at 1669 cm^{-1} showing the centre of the secondary structure of the amide I band spectrum

Collagen crosslinking is measured as changes in the amide I envelope [39, 46]. It was observed that the Raman band at $\sim 1669\text{ cm}^{-1}$ was present in the fractions containing the trivalent collagen cross-links whereas IR observed a band at 1632 cm^{-1} , but no band was evident at $\sim 1669\text{ cm}^{-1}$ [17, 27]. From literature studies [47, 48], biochemical analysis of collagen peptides showed that pyridinoline (Pyr) crosslinks result in a band at 1666 cm^{-1} . Therefore, the peak at $\sim 1669\text{ cm}^{-1}$ reflects pyridinoline cross-linked collagen peptides [29]. These observations from Raman and IR spectra provide additional information of changes in the amide I band. Most of the underlying bands of amide I arise from the structure of the collagen triple helix as well as the telopeptides (1632 , 1645 , 1655 , 1672 , and 1682 cm^{-1}). The intermolecular cross-linking of collagen is a key element in determining tensile strength and elasticity [49, 50].

For amide I, the Raman band area ratio of $1648/1681\text{ cm}^{-1}$ (non-reducible and reducible collagen types) was used for analyzing variations between loose and tight

leathers, whereas, for IR, the $1632/1650\text{ cm}^{-1}$ (triple helix and α -like helix collagen types) ratio was used, as shown in Fig. 6 (Additional file 1).

For quality assessment, a student t-test was carried out between the two ratio datasets. For Raman, the t-test gave a value of $p = 0.0008$, and for IR, $p = 6.8 \times 10^{-5}$. So, there are significant differences ($p < 0.05$) between the DA and OSP ratios, and, therefore, suitable to fit to a regression model.

3.2.2 Logistic regression

Quantitative classification of DA and OSP involves a continuous independent variable (peak area ratio) and a binary dependent variable (DA vs OSP), therefore a logistic regression (LR) algorithm [37] was devised to discriminate the samples using the SciKit Learn package [30] in Python 3.7.

A confusion matrix was generated from the output that describes the performance of classification. It summarises correct and incorrect spectra classification. It is useful for two-class classification and in measuring recall, precision and accuracy [18, 51, 52]. The confusion matrix for Raman and IR data obtained is presented in Tables 4 and 5.

The first entry in the confusion matrix is the number of correctly identified DA samples. i.e. $5/5$ which is a perfect classification whereas for OSP it is $4/5$ which is also close to perfect fit. Accuracy, precision and recall are of importance where:

$$\text{Accuracy} = \frac{TP + TN}{\text{Total}}$$

Table 3 Deconvoluted amide I band assignments for protein secondary structures

Wavenumber (cm^{-1})	Assignment
1615 ± 4	β -sheet
1645 ± 4	random coils
1653 ± 6	α -helix
1681 ± 3	β -turn
1698 ± 4	β -turn

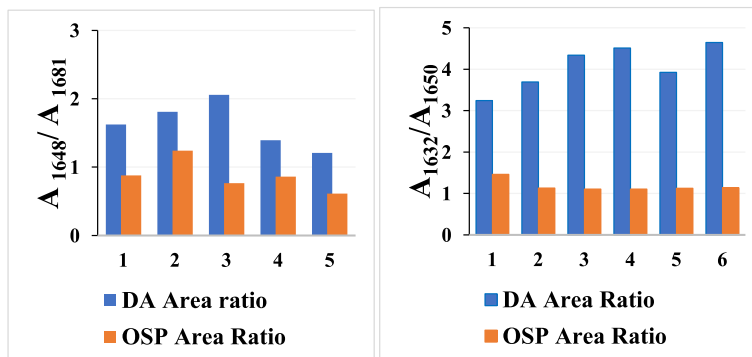


Fig. 6 Collagen crosslinking ratio shown as the curve fitted area ratio of Raman bands at 1648 cm⁻¹ to 1681 cm⁻¹ representing non-reducible and reducible collagen types respectively (left); IR bands at 1632 cm⁻¹ to 1650 cm⁻¹ represent the triple helix and α-like helix collagen types respectively (right)

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

where TP = true positive, TN = true negative, FP = false positive, and FN = false negative with DA arbitrarily set as True and OSP set as False.

The accuracy for the Raman data is 0.9 (90%), precision is 1.0 (100%), and the recall score is 0.8 (80%). The IR data presented in Table 4 shows the perfect classification of 6/6 from all six DA and OSP. This means all were correctly classified. The accuracy, precision and recall score is 1.0 (100%).

From the results obtained, it is evident that Raman peak area ratios and IR peak ratios are a good predictor in differentiating the leather type. Both techniques complement each other very well. There is a significant difference obtained in the recall score for Raman and IR data that provides the motivation for the multivariate analysis of the Raman data. Although univariate analysis is quite useful, it might be possible to still obtain a useful prediction from Raman spectra by using a multivariate analysis to reveal the differences, especially when there is a large dataset.

3.3 Multivariate analysis

Multivariate analysis can be used to quickly characterise the “types” or “classes” of spectra or samples present in a large data set. An unsupervised method, Principal

Components Analysis, is used that can determine the existence of classes in the data set without any assumptions of the number of classes. The classes are determined by transforming the data set, expressed in the original spectral variables, to a new description using variables (principal components) that maximise the separation between samples (the principal components are the eigenvectors of the variance-covariance matrix). A scores plot shows the samples plotted using the principal components. If distinct clusters of samples are observed in the scores plot, then classes exist in the data set. All spectral variables in the original data set have been used in the analysis presented here [27].

There is a supervised method, linear discriminant analysis, that assumes the existence of classes and then proceeds to constructs a function (the discriminant) that gives the best separation between the classes. LDA works on a similar approach to PCA, but LDA creates a linear function (the discriminator) that maximises the differences among the classes or groups [44]. It will show how well the classes are separated as well as where the classification fit is robust and where it is misinterpreted. To demonstrate the best performance of classification in a robust model, combinations of both PCA and LDA were attempted. A potential issue with LDA is that it will *always* sort samples into classes, so it is difficult to determine if the model contains errors. However, performing PCA prior to LDA can independently confirm the existence of classes in the data set. The principal components from the PCA analysis can also be used to construct the discriminant function in LDA (PCA-LDA).

Table 4 Confusion matrix for Raman univariate analysis

Actual Values	Classifier Prediction	
	DA	OSP
DA	5	1
OSP	0	4

Table 5 Confusion matrix for IR univariate analysis

Actual Values	Classifier Prediction	
	DA	OSP
DA	6	1
OSP	0	6

Figure 7 below shows the loading plots of first three principal components. PC1 explains 59.1% of the data while PC2 and PC3 explain 16.2% and 11.7%, respectively.

The loading plots shown in Fig. 7 indicate which spectral bands contribute most to the variance described by the principal component. The OSP average spectrum is used as a reference for comparison of loadings. This gives an understanding of the origin of differences between the samples corresponding to spectral variations. The strong contribution in PC1 and PC2 is from the C=O stretch around 1669 cm^{-1} which is usually from amide I band and is mainly proteins and lipids. PC2 has spectral contributions from the Amide III band, around 1243 cm^{-1} , which is purely collagen and from the CH_2 wag with a broad noisy band around 1340 cm^{-1} which indicated collagen and lipids. PC3 shows another significant contribution around 1100 cm^{-1} that is broadly from C-O-C modes, which is mainly protein [33, 50]. So, the

loadings are showing that DA and OSP samples are differentiated due to their protein and lipid content. These observations are consistent with the identification of the tight and loose marker bands as discussed in Fig. 4. It appears that these bands are responsible for the classification of the samples using Raman and there was a distinguished difference between the two leather types.

Figure 8(a-c) shows the two-dimensional (2-D) score plots, with 95% confidence ellipse, of combination of two principal components with the aim to find the separation between DA and OSP. The screeplot in Fig. 8d shows the proportion of the percentage of variance that is accounted for by the principal components.

DA and OSP were not separable as clusters within principal components, but OSP samples show a significant separation along PC3 and DA samples are on the positive side of PC2. No perfect distinction is found between the replicates of OSP and DA in other

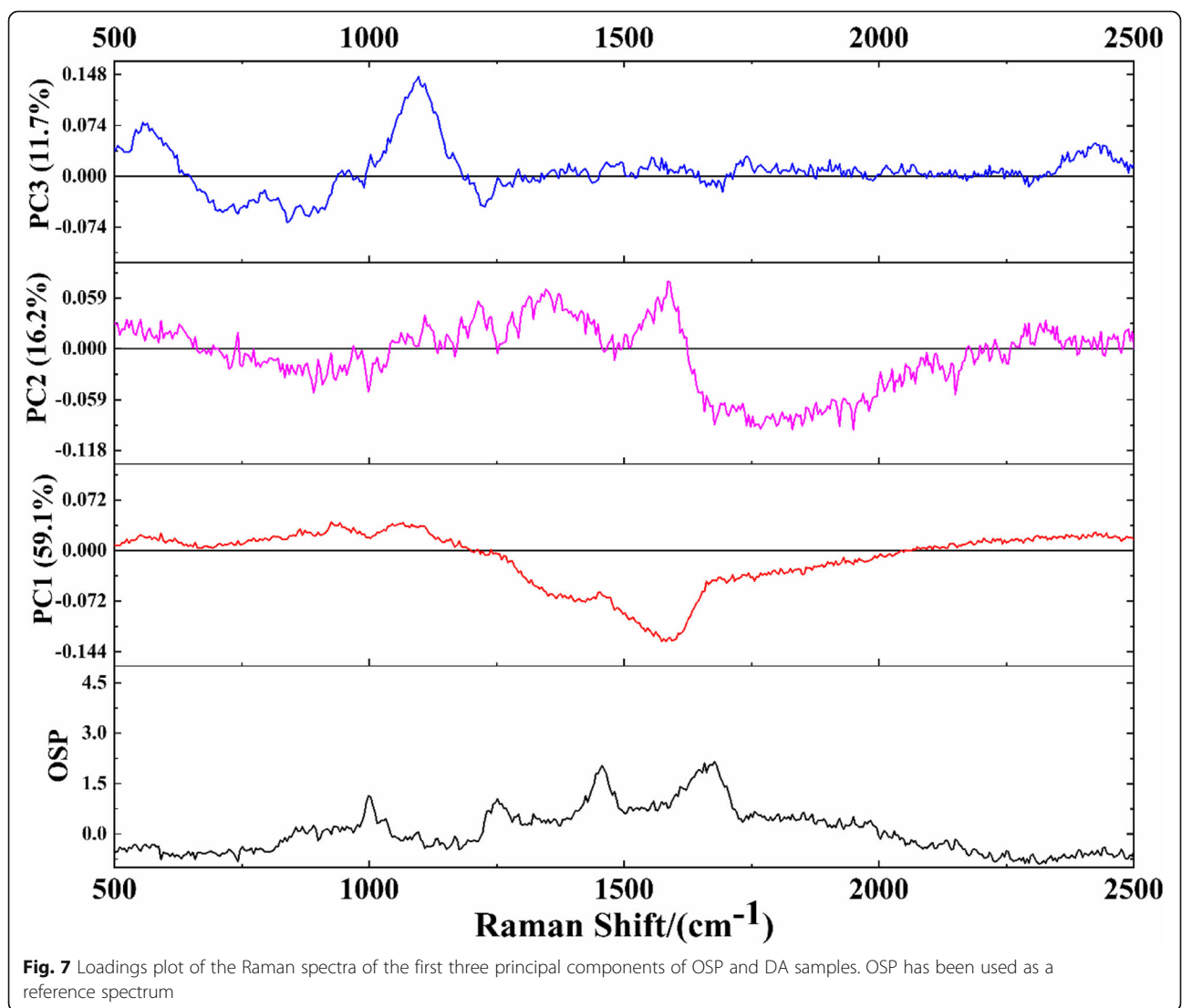
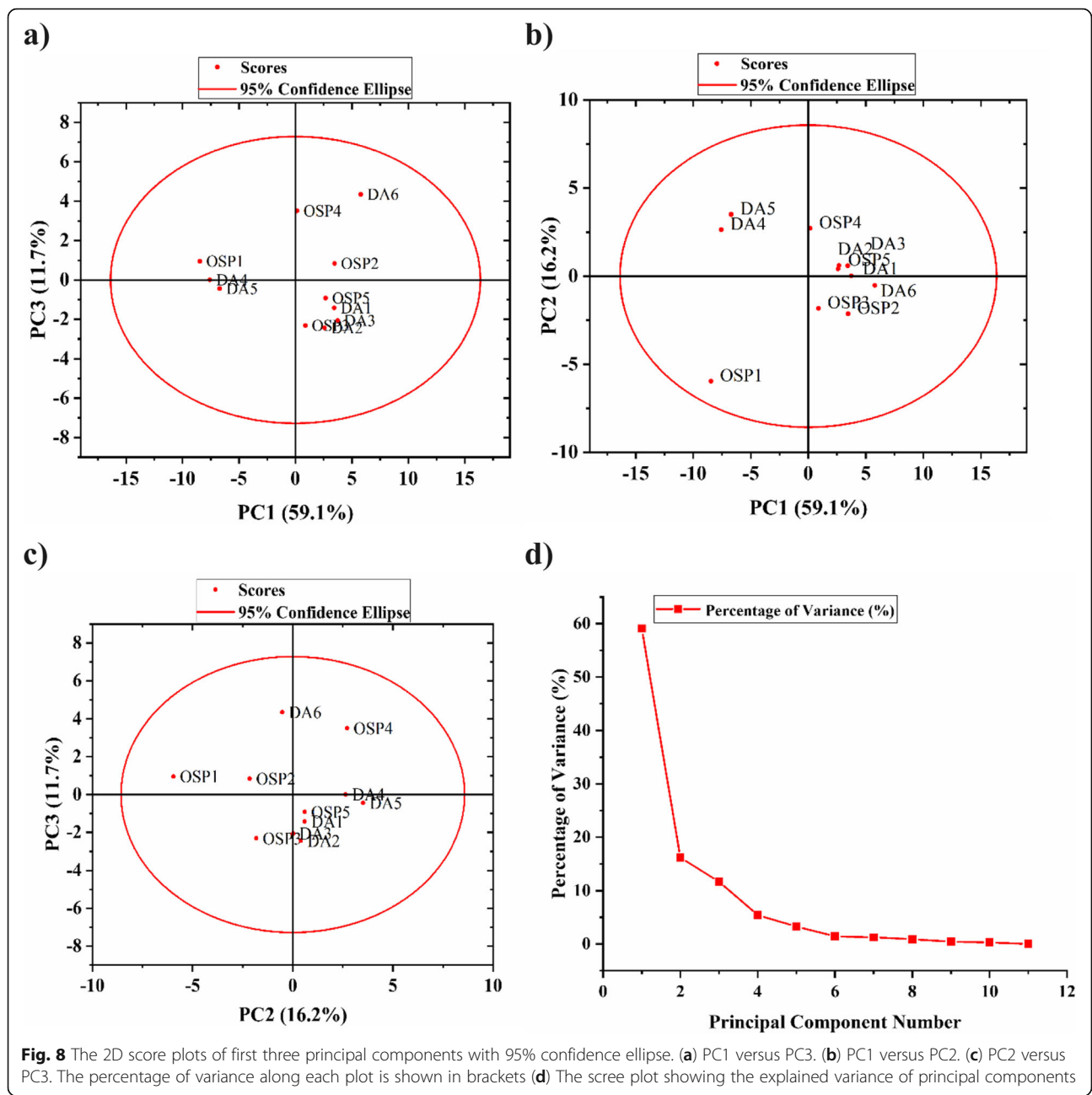


Fig. 7 Loadings plot of the Raman spectra of the first three principal components of OSP and DA samples. OSP has been used as a reference spectrum

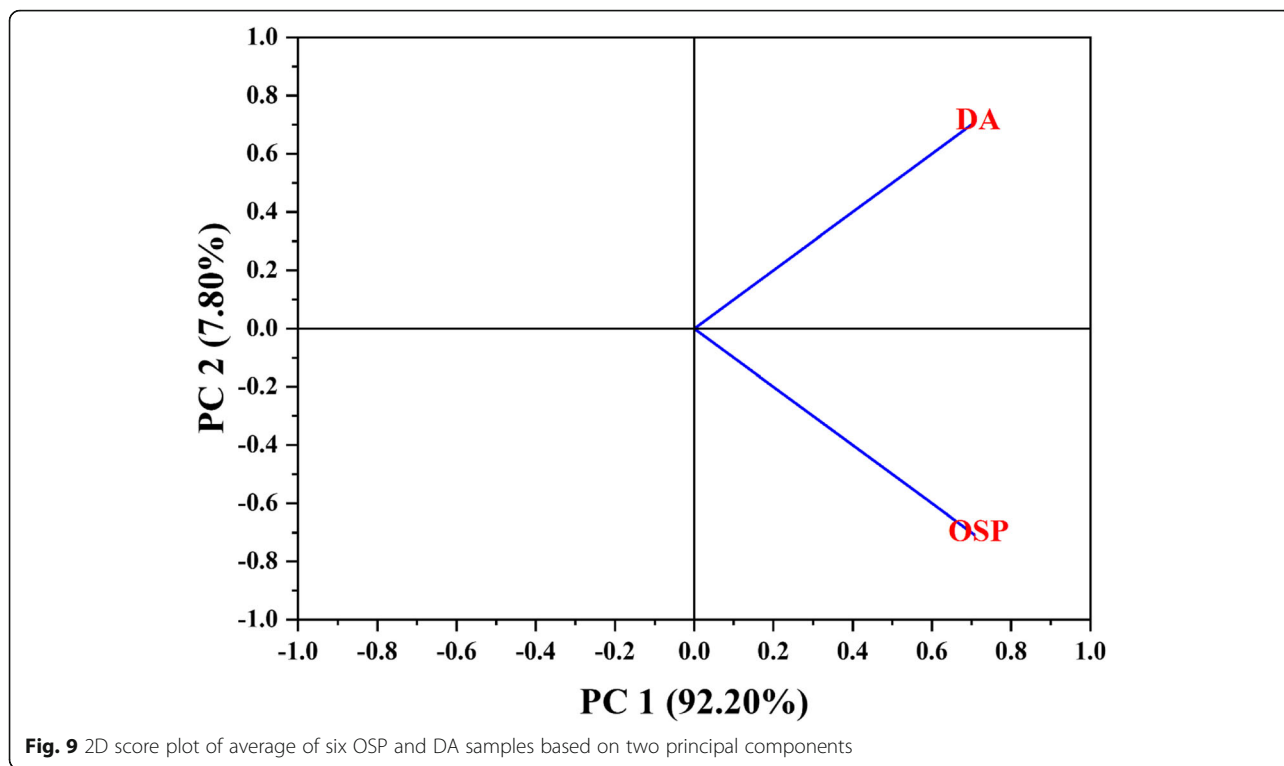


dimensions. A few points of OSP and DA samples in Fig. 8b overlap with each other. PCA scores plots reveal the intra and intergroup variation between loose and tight wet blue samples [27, 46].

After comparing the sample replicates at individual level, principal component analysis was performed on the average spectra of loose and tight replicates. An interesting observation shown in Fig. 9 is that OSP and DA are symmetrically grouped at the positive and negative extremes which is contributed by PC2, which is 7.8% and discriminates both, respectively. This reduced dataset formed after averaging six

samples of OSP and DA makes the differentiation more identifiable and provides a significant role in visualising it clearly.

By displaying the data along the directions of maximal variance, PCA analysis demonstrates that Raman spectra of Wet Blue can be separated into classes (i.e. loose and tight). However, the principal components might not give the maximum separation between the classes. Linear discriminant analysis was used to construct a function (the discriminant function) that maximised the class separability. LDA assumed that the data was Gaussian distributed, that all rows must belong to one group



(samples are mutually exclusive) and that the variances are the same for both groups.

The original variables, or the principal components can be used to construct the discriminant function, the principal components are used as they have the advantage of being independent. If two principal components are used (so the scores plot is planar), the

LDA process finds the line in the scores plot plane that maximises separability.

When LDA is done on the PC scores, the mean centre of each grouping is calculated, and each spectrum is predicted to belong to one of the groups based on its distance from the centre of the group. The accuracy of the prediction is an indication of how well the groups are separated [15, 23].

A leave-one-out cross validation method [53] was utilised to train the LDA classifier where one sample is left out of the calibration model and predicted with the results obtained. Results are plotted in Fig. 10, which shows the observed group with the predicted group along with the cross-validation summary (Table 6). It also shows that one of the six DA samples was falsely identified as OSP and two of the five OSP samples were

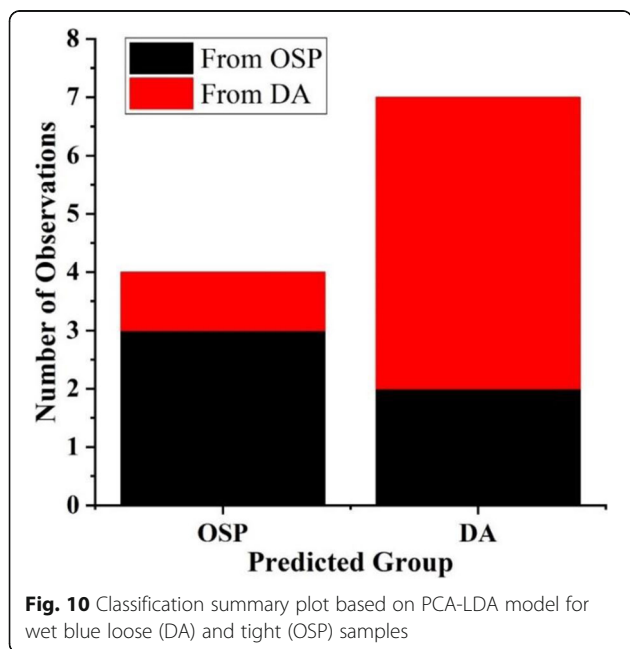


Table 6 Confusion matrix for the classification of OSP and DA based on the PCA-LDA model

Predicted Groups			
Groups	OSP	DA	Total
OSP	3	2	5
	60%	40%	100%
DA	1	5	6
	16.67%	83.33%	100.00%
Total	4	7	11
	36.36%	63.64%	100.00%

falsely identified as DA. The error rate for cross validation of the training data is 12.33%.

The Wilk's Lambda test was conducted on the discriminant variable and found that the discriminant function is highly significant ($p < 0.05$) in agreement with the classification summary.

The cross-validation summary table shows that OSP has a classification accuracy of 60% and DA has 83.33% which proves that both are mutually exclusive as was expected.

4 Conclusion

In summary, the work presented here used Raman and Infrared spectroscopy to investigate the variations in loose and tight wet blue leather. To the best of our knowledge, this is the first study done in depth using ratiometric and chemometric analysis to identify and quantify the difference between two wet blue samples of OSP and DA. Vibrational spectroscopy with advanced spectral analysis can quantify the biomolecules which impact the quality, strength and sustainability of leather.

Classification from the peak area ratios was done using logistic regression that gives 100% accuracy for IR data and 90% accuracy for Raman data. Multivariate analysis has supported the Raman results for OSP and DA in describing the difference between the groups providing a clear representation of underlying biological differences. This study is a proof of principle to employ vibrational spectroscopy for quality assessment of leather.

Identification of issues at the raw skin stage and differentiating the changes occurring at each stage of leather processing will be the next area of further work so that only high-quality leather can be obtained with no defects. The analysis of Raman spectra in this work classifies leather samples based mostly on their chemical composition as this factor has the strong influence on the shape of the Raman spectra. Structural factors are also likely to be important. Polarised Raman microscopy can provide structural information that complements the chemical information acquired from the spectral data alone. Further research will analyse the Amide I and Amide III bands using polarised Raman microscopy to provide information on cross-linking between microstructures in the samples.

5 Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s42825-019-0017-5>.

Additional file 1. Supporting Information.

Abbreviations

ATR – FTIR: Attenuated Total Reflectance - Fourier Transform InfraRed;
DA: Distal axilla; LDA: Linear discriminant analysis; LR: Logistic regression;

OSP: Official sampling position; PCA: Principal component analysis;
Pyr: Pyridinoline

Acknowledgements

This work was supported by NZ Leather and Shoe Research Association (LASRA*), Palmerston North, New Zealand through the Ministry of Business, Innovation and Employment (MBIE) grant number LSRX1801. We would like to thank Massey University, Palmerston North, New Zealand for accessing Raman Laboratory for acquiring the data. Wenkai Zhang for the help in the sample preparation.

Authors' contributions

M.M. conducted all Raman and InfraRed experiments, acquiring the data, analysis and writing the manuscript. MW scripted the Python algorithm for pre-processing and logistic regression. R.N. had the input into targeting wet blue. R.N., C.M. and G. H. for the sample collection and preparing wet blue. All authors read and approved the final manuscript.

Funding

The research is funded by the Ministry of Business, Innovation and Employment (MBIE) grant number LSRX1801.

Availability of data and materials

Curve-fitted area ratios for Raman and IR spectra are listed in the supporting information.

Competing interests

The authors declare that there are no competing interests.

Author details

¹NZ Leather and Shoe Research Association (LASRA*), Palmerston North, New Zealand. ²School of Fundamental Sciences, Massey University, Palmerston North, New Zealand.

Received: 27 June 2019 Accepted: 26 November 2019

Published online: 05 February 2020

References

- Carter K. Don't Hide from the truth.2008.<https://www.theguardian.com/lifeandstyle/2008/aug/27/ethicalfashion>, Accessed 27 Aug 2008.
- Masciana P. World statistical compendium for raw hides and skins, leather and leather footwear 1998-2014. Rome: FAO; 2015. p. 1–122.
- Covington AD. Tanning chemistry: the science of leather. Cambridge: Royal Society of Chemistry; 2009.
- Wells HC, Holmes G, Haverkamp RG. Looseness in bovine leather: microstructural characterization. *J Sci Food Agric*. 2016;96(8):2731–6.
- Shelly DC. Raman spectroscopy of wet blue bovine leather. *J Am Leather Chem Assoc*. 2009;104:9-15.
- Prabakar S, Whitby CP, Henning AM, Holmes G. The effect of Cloisite® Na+ nanoclay filler on the morphology and mechanical properties of loose leather. *J Am Leather Chem Assoc*. 2016;111(5):2.
- Zhang Y, Ingham B, Cheong S, Ariotti N, Tilley RD, Naffa R, Holmes G, Clarke DJ, Prabakar S. Real-time synchrotron small-angle X-ray scattering studies of collagen structure during leather processing. *Ind Eng Chem Res*. 2017;57(1):63–9.
- Zhang Y, Mansel BW, Naffa R, Cheong S, Yao Y, Holmes G, Chen HL, Prabakar S. Revealing molecular level indicators of collagen stability: minimizing chrome usage in leather processing. *ACS Sustain Chem Eng*. 2018;6(5):7096–104.
- Van Stiphout T, Pel L, Galvosas P, Prabakar S, Holmes G. NMR transverse relaxation analysis of leather looseness; 2015.
- Matousek P, Stone N. Emerging concepts in deep Raman spectroscopy of biological tissue. *Analyst*. 2009;134(6):1058–66.
- Wachsmann-Hogiu S, Weeks T, Huser T. Chemical analysis in vivo and in vitro by Raman spectroscopy—from single cells to humans. *Curr Opin Biotech*. 2009;20(1):63–73.
- Larkin P. Infrared and Raman spectroscopy: principles and spectral interpretation. Amsterdam: Elsevier; 2017.
- Chen Y, Dai J, Zhou X, Liu Y, Zhang W, Peng G. Raman spectroscopy analysis of the biochemical characteristics of molecules associated with the malignant transformation of gastric mucosa. *PLoS One*. 2014;9(4):e93906.

14. Suci PA, Vrani JD, Mittelman MW. Investigation of interactions between antimicrobial agents and bacterial biofilms using attenuated total reflection Fourier transform infrared spectroscopy. *Biomaterials*. 1998;19(4–5):327–39.
15. Ahmed R, Wang W, Zia AW, Lau C. Collagen formation observed from healing calvarial defects with principal component analysis of Raman scattering. *Analyst*. 2018;143(19):4614–22.
16. Unal M, Jung H, Akkus O. Novel Raman spectroscopic biomarkers indicate that postyield damage denatures bone's collagen. *J Bone Miner Res*. 2016; 31(5):1015–25.
17. Kumar S, Verma T, Mukherjee R, Ariese F, Somasundaram K, Umapathy S. Raman and infra-red microspectroscopy: towards quantitative evaluation for clinical research by ratiometric analysis. *Chem Soc Rev*. 2016;45(7):1879–900.
18. Eberhardt K, Matthäus C, Marthandan S, Diekmann S, Popp J. Raman and infrared spectroscopy reveal that proliferating and quiescent human fibroblast cells age by biochemically similar but not identical processes. *PLoS One*. 2018;13(12):e0207380.
19. Yamaguchi Y, Haginaka J, Kunitomo M, Yasuda H, Bandō Y. High-performance liquid chromatographic determination of desmosine and isodesmosine in tissues and its application to studies of alteration of elastin induced by atherosclerosis. *J Chromatogr B Biomed Sci Appl*. 1987;422:53–9.
20. Naffa R, Holmes G, Ahn M, Harding D, Norris G. Liquid chromatography-electrospray ionization mass spectrometry for the simultaneous quantitation of collagen and elastin crosslinks. *J Chromatogr A*. 2016;1478:60–7.
21. Naffa R, Watanabe S, Zhang W, Maiment C, Singh P, Chamber P, et al. Rapid analysis of pyridinoline and deoxypyridinoline in biological samples by liquid chromatography with mass spectrometry and a silica hydride column. *J Sep Sci*. 2019;42(8):1482–8.
22. Naffa R, Maiment C, Ahn M, Ingham B, Hinkley S, Norris G. Molecular and structural insights into skin collagen reveals several factors that influence its architecture. *Int J Biol Macromol*. 2019;128:509–20.
23. Carden A, Rajachar R, Morris MD, Kohn D. Ultrastructural changes accompanying the mechanical deformation of bone tissue: a Raman imaging study. *Calcif Tissue Int*. 2003;72(2):166–75.
24. Timchenko E, Zherdeva L, Timchenko P, Volova L, Ponomareva U. Detailed analysis of the structural changes of bone matrix during the demineralization process using Raman spectroscopy. *Phys Procedia*. 2015;73:221–7.
25. McCreadie BR, Morris MD, Chen T-c, Rao DS, Finney WF, Widjaja E, et al. bone tissue compositional differences in women with and without osteoporotic fracture. *Bone*. 2006;39(6):1190–5.
26. Ahmed R, Lau C, editors. Bone mineralization and collagen formation observed with principal component analysis of Raman scattering from healing calvarial defects. In: Proceedings of the SPIE BIOS conference of Imaging, Manipulation, and Analysis of Biomolecules, Cells, and Tissues XVII; 2019. San Francisco: International Society for Optics and Photonics; 2019. p. 10881.
27. Alebrahim MA. ATR-FTIR and Raman imaging to study permanent and primary teeth from different places and ages [dissertation on the internet]. Institute of photonic technology, Jena university; Accessed 8 Mar 2013. Available from: <http://nbn-resolving.org/urn:nbn:de:gbv:27-20130704-101500-0>.
28. Jaafreh S, Breuch R, Günther K, Kreyenschmidt J, Kaul P. Rapid poultry spoilage evaluation using portable fiber-optic Raman spectrometer. *Food Anal Methods*. 2018;11(8):2320–8.
29. Lin S-Y, Li M-J, Cheng W-T. FT-IR and Raman vibrational microspectroscopies used for spectral biondiagnosis of human tissues. *J Spectrosc*. 2007;21(1):1–30.
30. Buitinck L, Louppe G, Blondel M, Pedregosa F, Mueller A, Grisel O, et al. API design for machine learning software: experiences from the SciKit-Learn project. <https://arxiv.org/abs/1309.0238>. 2013.
31. Ahmed R, Law AWL, Cheung TW, Lau C. Raman spectroscopy of bone composition during healing of subcritical calvarial defects. *Biomed Opt Express*. 2018;9(4):1704–16.
32. Chen Z, Butke R, Miller B, Hitchcock CL, Allen HC, Pivoski SP, et al. Infrared metrics for fixation-free liver tumor detection. *J Phys Chem B*. 2013;117(41): 12442–50.
33. Masic A, Bertinetti L, Schuetz R, Galvis L, Timofeeva N, Dunlop JW, et al. Observations of multiscale, stress-induced changes of collagen orientation in tendon by polarized Raman spectroscopy. *Biomacromolecules*. 2011; 12(11):3989–96.
34. Qin L, Bi J-R, Li D-M, Dong M, Zhao Z-Y, Dong X-P, et al. Unfolding/refolding study on collagen from sea cucumber based on 2D fourier transform infrared spectroscopy. *Molecules*. 2016;21(11):1546.
35. Depciuch J, Kaznowska E, Zawlik I, Wojnarowska R, Cholewa M, Heraud P, et al. Application of Raman spectroscopy and infrared spectroscopy in the identification of breast cancer. *Appl Spectrosc*. 2016;70(2):251–63.
36. Huang Z, McWILLIAMS A, Lam S, English J, McLEAN DI, Lui H, et al. Effect of formalin fixation on the near-infrared Raman spectroscopy of normal and cancerous human bronchial tissues. *Int J Clin Oncol*. 2003;23(3):649–55.
37. Davidson B, Murray AA, Elfick A, Spears N. Raman micro-spectroscopy can be used to investigate the developmental stage of the mouse oocyte. *PLoS One*. 2013;8(7):e67972.
38. Bhushan B, Pradhan A. Detection of breast & cervical cancer using Raman spectroscopy. *IJRSET*. 2013;2(8):3420–5.
39. Neuman RE, Logan MA. The determination of collagen and elastin in tissues. *J Biol Chem*. 1950;186(2):549–56.
40. Movasaghi Z, Rehman S, Rehman IU. Raman spectroscopy of biological tissues. *Appl Spectrosc Rev*. 2007;42(5):493–541.
41. Schulz H, Baranska M. Identification and quantification of valuable plant substances by IR and Raman spectroscopy. *Vib Spectrosc*. 2007;43(1):13–25.
42. Yu P. An emerging method for rapid characterization of feed structures and feed component matrix at a cellular level and relation to feed quality and nutritive value. *Arch Anim Nutr*. 2006;60(3):229–44.
43. Saguer E, Alvarez P, Sedman J, Ramaswamy H, Ismail A. Heat-induced gel formation of plasma proteins: new insights by FTIR 2D correlation spectroscopy. *Food Hydrocoll*. 2009;23(3):874–9.
44. Ur Rehman I, Movasaghi Z, Rehman S. *Vibrational spectroscopy for tissue analysis*: CRC press; 2012.
45. Hanson DA, Eyre DR. Molecular site specificity of pyridinoline and pyrrole cross-links in type I collagen of human bone. *J Biol Chem*. 1996;271(43): 26508–16.
46. Khalid M, Bora T, Al Ghaithi A, Thukral S, Dutta J. Raman spectroscopy detects changes in bone mineral quality and collagen cross-linkage in staphylococcus infected human bone. *Sci Rep*. 2018;8(1):9417.
47. Fox SA. Raman micro-spectroscopy and multivariate analysis for the. *Skin*; 2013.
48. Gamsjaeger S, Robins SP, Tatakis DN, Klaushofer K, Paschalis EP. Identification of pyridinoline trivalent collagen cross-links by Raman microspectroscopy. *Calcif Tissue Int*. 2017;100(6):565–74.
49. Bailey AJ. The chemistry of collagen cross-links and their role in meat texture. *Proc Recip Meat Conf*. 1989;42:127–35.
50. Ami D, Natalello A, Mereghetti P, Neri T, Zanoni M, Monti M, et al. FT-IR spectroscopy supported by PCA-LDA analysis for the study of embryonic stem cell differentiation. *J Spectrosc*. 2010;24(1–2):89–97.
51. Ramos I. The potential of Raman spectroscopy for cytological diagnosis of ThinPrep® samples from a cervical Cancer screening population. Doctoral thesis, DIT; 2017.
52. Renugopalakrishnan V, Carreira L, Collette T, Dobbs J, Chandraksasan G, Lord R. Non-uniform triple helical structure in chick skin type I collagen on thermal denaturation: Raman spectroscopic study. *Z Naturforsch C*. 1998; 53(5–6):383–8.
53. Visa S, Ramsay B, Ralescu AL, Van Der Knaap E. Confusion matrix-based feature selection. *MAICS*. 2011;710:120–7.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.