# Pocket-size near-IR spectrometers for rapid examination of contaminated textile fibres at the crime-scene

Hamad S. Rashed [a,b], Andrew J. Parrott [c], Alison Nordon [c], Matthew J. Baker [a,e], David S. Palmer [d,*]

[a] *Department of Pure and Applied Chemistry, University of Strathclyde, Technology and Innovation Centre, 99 George Street, Glasgow G1 1RD, UK*
[b] *Trace Evidence Section, General Department of Forensic Science and Criminology, Dubai Police GHQ, Dubai, United Arab Emirates*
[c] *Department of Pure and Applied Chemistry and Centre for Process Analytics and Control Technology, University of Strathclyde, Royal College Building, 204 George Street, Glasgow G1 1XW, UK*
[d] *Department of Pure and Applied Chemistry, University of Strathclyde, Thomas Graham Building, 295 Cathedral Street, Glasgow G1 1XL, UK*
[e] *School of Medicine, Faculty of Clinical and Biomedical Sciences, University of Central Lancashire, Fylde Rd, Preston, PR1 2HE*

## ARTICLE INFO

## ABSTRACT

In a typical forensic investigation, fabric analysis plays a vital role in solving different crimes. Several types of textile fibre materials (cotton, polyester, denim, polypropylene, polycotton, and viscose) were analysed in the presence of common contaminants (blood, rainwater, seawater, sand and gunshot-residue) to evaluate the performance of two NIR spectrometers for in situ analysis of different crime scene conditions. The spectrometers used were SCIO® by Consumer Physics and NIRscan Nano by Texas Instruments. The SCIO instrument covers the third overtone region (740–1070 nm) and NIRscan Nano instrument covers the first and second overtone regions (900–1700 nm). Spectra from both instruments were pre-processed using the PRFFECTv2 software to eliminate noise and smooth the data for classification model construction. The models showed high accuracy, sensitivity and specificity with a range of 69–100% for binary classification (one class versus others) and range of 76–100% for multi-class classification of fibre material. This study shows for the first time the capability of pocket-size spectrometers coupled with random forest models for classification of fibre material in the presence of common contaminants in a rapid and non-destructive manner, and so is suitable for investigation of crime scenes.

## 1. Introduction

In a typical forensic investigation, fabric analysis plays a vital role in providing clues and leads. Traditionally, fibre analysis involves a microscopic comparison between known and unknown fibre samples [1]. This is a slow and tedious task which requires a well-equipped laboratory and trained staff to carry out the whole process. Historically, identification of specific fibre types has provided the key evidence in many different types of criminal cases including murder, sexual assault, and hit and run. The fibre materials most frequently involved in criminal cases are cotton, polyester, polypropylene, polycotton, viscose, and denim [2]. Rapid, efficient and precise methods for fibre confirmation at the crime scene are highly sought to enable the police, forensic and judicial services to benefit from better data sooner.

In addition to the identification of the fibre, the identification of the dye used within the textile is often of interest [3]. Commonly, in the forensic examination of dyed fibres, extraction methods are used and combined with chromatographic or electrophoretic analysis for dye characterisation [4,5]. High performance liquid chromatography (HPLC) has been widely applied for dye analysis of fibres in forensic science in cases of single and mixed dyes [6]. However, this method requires destruction of the forensic sample which is not considered favourable during the forensic process. Alternatively, near-infrared spectroscopy (NIRS) is a technique which does not contaminate or damage the fabric sample during the forensic analysis process and is considered to be more advanced and efficient. In NIRS, light is absorbed at specific frequencies corresponding to molecular overtones and combination vibrations [7].

NIRS has shown promising results in different areas such as the food industry, polymer analysis, fuel analysis and textile fibres identification [8]. A few studies have explored the use of NIRS coupled with the random forest algorithm in different areas, as it has been shown to be a
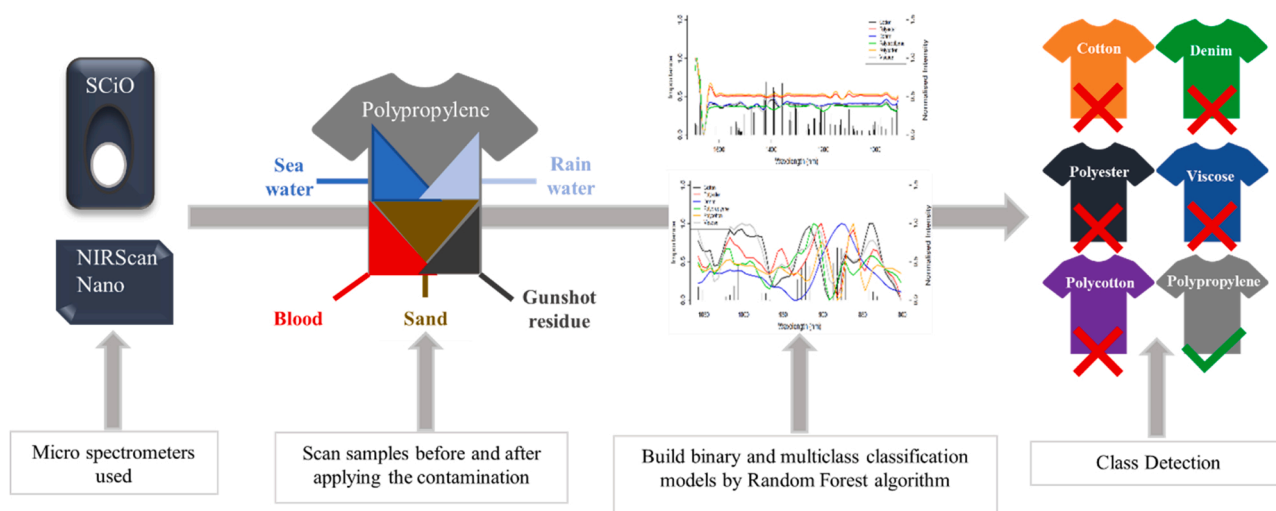
---

**Fig. 1.** Workflow for in situ analysis of fabric samples using SCIO and NIRScan Nano sensors: application of the contaminant to the fabric samples; spectral acquisition; construction of models; and classification of fabric type.

promising algorithm for classification purposes [9-11]. NIRS has been investigated for forensic examination of different materials such as illicit drugs, faked documents, currency, bloodstains and explosive materials [11]. Benchtop NIRS instruments and cameras have been widely investigated in the identification of the natural materials that make up textile fibres, including those from plants and animals [12-14]. Several studies have shown NIRS joined with principal component analysis (PCA), partial least squares discriminant analysis (PLS-DA) and soft-independent modelling of class analogies (SIMCA) can be used to discriminate between textile types with accuracies, sensitivities and specificities above 90% [13,15]. Such classification studies have been applied to identify animal fibres [16], cashmeres [17], wool-cashmere blends [18], cotton fibres [19] and other fibres from different plant origins [13,20]. Moreover, accurate classification of four fibre materials of plant and animal origin (cotton, flax, silk and cashmere) has been achieved by visible-NIRS combined with PCA and least squares support vector machines (PCA-LS-SVM) [15].

While some previous studies have shown satisfactory results for fabric analysis, all of these approaches are laboratory based and time consuming. None of these approaches are applicable for in situ analysis at a crime scene. The optics industry has recently been revolutionised by the commercialisation of miniaturised NIR sensors, and these have begun to be used for non-invasive forensic examination. This has included the use of miniature NIR instruments for differentiation of authentic and counterfeit banknotes [14], and instant identification of body fluid stains (blood) at the crime scene [21]. Applications for textile analysis are also rising as miniature NIR instruments are low-cost, compact, end-user friendly, small in size and have decent performance [21-23]. In a previous study, we investigated the use of the pocket-size NIR spectrometers (SCIO and NIRscan Nano) for identification of clean textile fibres in-situ but that work did not consider the kind of fibre contamination that would be expected at crime scenes [22].

The aim of the current study is to investigate NIR nano size spectrometers (SCIO and NIRscan Nano) for identifying and classifying contaminated textile fibres for potential applications in the crime scene. Both instruments are non-destructive, record the diffuse reflectance of the sample, and have short scanning times (few seconds). Moreover, no sample preparation is needed for fibres (i.e. dye extraction), which make the instruments suitable for direct analysis of fibres at the crime scene. The SCIO covers the wavelength range of 740 – 1070 nm, which represents the third overtone in the NIR spectrum [21]. The NIRscan Nano comprises the wavelength range of 900 – 1700 nm, corresponding to the first and second overtone regions [24]. The data were further refined

using a variety of spectral pre-processing techniques from which the best ones were identified. Classification was performed with random forest [25,26].

## 2. Materials and methods

### 2.1. Materials

In this study, the textile fibres classes were the common fibre materials found at different crime scenes worldwide. The fabrics were contaminated with different substances that they might be commonly exposed to at the crime scene: blood; gunshot-residue (GSR); sand; rain water; and sea water. A total of 210 fabric samples were collected from various British and European sources. These comprised 35 samples of six different materials that are commonly encountered in forensic work: cotton, polyester, denim, polypropylene, polycotton and viscose.

The contaminants were applied to all samples in each fabric classes in a controlled environment in the laboratory. A defibrinated horse blood was applied as a 50 µL pipette droplet on to the fabric surface. For the gunshot-residue, the fabrics were hung on a board, and were shot with a 9 mm shotgun, in an indoor shooting range. For the other contamination types, the fabrics were immersed into vessels of sand, rainwater or sea water as appropriate.

### 2.2. Data collection

Spectra of textile fabrics were collected using the NIRscan Nano and SCIO spectrometers. The materials were placed on a flat benchtop, and a shade was used and the curtains were closed for measurements made using the SCIO and NIRscan Nano instruments, respectively, to avoid any external light source.

The distance from the SCIO instrument to the fabric sample was 5 mm for all measurements. Spectra were collected over the 740 – 1070 nm wavelength range using the SCIO Lab mobile App for iOS and data were stored in the cloud. The scan resolution for the SCIO instrument is 5 – 10 nm and the illumination spot diameter is 1 cm. The NIRscan Nano sensor was placed directly in contact with the samples, as this was found to be the optimal position after some trials. The NIRscan Nano sensor was used to acquire spectra over the 900 – 1700 nm wavelength range with a resolution of 6.35 nm and was controlled by PC software (NIRscan Nano EVM). A labsphere Spectralon material 99% standard disc was used as a reference sample.

Spectra were acquired of five different areas for each of the 210

**Table 1**

The combinations of pre-processing methods used in the building of binary and multiclass classification random forest models in PRFFECTv2 with the SCIO and NIRscan Nano datasets.

| Model | Pre-processing method | | | |
|---|---|---|---|---|
| | Binning | Smoothing | Normalisation | Baseline Correction |
| **Binary classification of SCIO** | | | | |
| 1 | 5 | SG filtering (order 1) | Vector | Second Derivative |
| 2 | 5 | Local polynomial fit with Gaussian width 2 | Vector | Second Derivative |
| 3 | 4 | SG filtering (order 1) | Vector | Second Derivative |
| 4 | 4 | Local polynomial fit with Gaussian width 4 | Vector | Second Derivative |
| 5 | 5 | 0 | min/max | Second Derivative |
| **Binary classification of NIRScan Nano** | | | | |
| 1 | 2 | 0 | Vector | Second Derivative |
| 2 | 4 | 0 | Vector | Second Derivative |
| 3 | 0 | 0 | min/max | Second Derivative |
| 4 | 2 | 0 | min/max | Second Derivative |
| 5 | 5 | 0 | min/max | Second Derivative |
| **Multiclass classification of SCIO** | | | | |
| 1 | 5 | SG filtering (order 1) | Vector | Second Derivative |
| 2 | 5 | Local polynomial fit with Gaussian width 2 | Vector | Second Derivative |
| 3 | 0 | SG filtering (order 1) | Vector | Second Derivative |
| 4 | 5 | SG filtering (order 2) | min/max | Second Derivative |
| 5 | 5 | SG filtering (order 1) | auto-scaling | Second Derivative |
| **Multiclass classification of NIRScan Nano** | | | | |
| 1 | 5 | Local polynomial fit with Gaussian width 2 | Vector | Second Derivative |
| 2 | 2 | Wavelet denoising -filter length of 2 | Vector | Second Derivative |
| 3 | 0 | Wavelet denoising -filter length of 4 | Vector | Second Derivative |
| 4 | 0 | Wavelet denoising -filter length of 4 | min/max | Second Derivative |
| 5 | 4 | Wavelet denoising -filter length of 8 | min/max | Second Derivative |

Additionally, the RF model uses the Gini impurity metric to identify the important peaks in the spectra.

fabric samples using the two NIR sensors, which gave rise to a total of 1050 spectra for each sensor. A sampling area of 10 cm$^2$ was chosen to avoid scanning the same area twice. The spectral acquisition time was less than 5 s for both instruments.

The fabric sample data were categorised first by material, which is the mandatory attribute for the identification of each fabric, then colour and sample ID. These data categories were chosen as important meta-data for further modelling and validation. Fig. 1 summarises the data collection and analysis workflow of both spectrometers.

*2.3. Data pre-processing and analysis*

Data from the SCIO and NIRscan Nano instruments were pre-processed, and classification models were constructed using PRFFECTv2, which is a user-friendly software package written for the R Statistical Computing Environment. It was initially released for spectral analysis using RF classification but has since been extended to incorporate over 100 different classification algorithms and many different approaches for spectral pre-processing.

To build and validate the machine learning models, the data were subdivided 70%:30% to form the training and test sets, respectively, based on sample ID so that spectra from a single fabric sample only appeared in one partition of the data. RF hyperparameters were selected by a grid search to maximise Cohen's Kappa statistics computed from 5-fold cross-validation on the training data. The best selected model was then used to predict the data in the test set. The complete process was repeated 51 times using different random seeds to partition the training and test sets. The statistics reported below are averages over all 51 resamples. This process corresponds to nested cross-validation in which 5-fold cross-validation in the inner loop is used to tune the machine learning hyperparameters, and 51-fold Monte Carlo cross-validation in the external loop is used as an independent test of the model's predictive accuracy. Both binary (one material versus all others) and multiclass classification models were built using random forest. A total of six binary classification models were built, one for each of the different fibre types. In each case, the fibre of interest was treated as the positive class (e.g. cotton) and the union of all other fibre types formed a single negative class (e.g. polyester, polycotton, denim, viscose, and polypropylene).

PRFFECTv2 calculates several different metrics to assess the performance of the machine learning models. *Accuracy* is the fraction of the spectra that are classified correctly. *Sensitivity* is the fraction of the real positive class that are classified correctly. *Specificity* is the fraction of the real negative class that are classified correctly. Each of these statistics is computed for both cross-validated (e.g. $Accuracy_{CV}$) and test set ($Accuracy_{TE}$) predictions. The statistics are defined below in terms of the number of real positives (P), real negatives (N), true positives (TP), true negatives (TN), false positives (FP) and false-negative (FN) [23,27].

$$Sensitivity = \frac{TP}{(TP + FN)} = \frac{TP}{P}$$

$$Specificity = \frac{TN}{(TN + FP)} = \frac{TN}{N}$$

$$Accuracy = \frac{(TP + TN)}{(TP + FP + FN + TN)} = \frac{(TP + TN)}{(P + N)}$$

PRFFECTv2 provides several different data pre-processing methods, including binning with various factors, smoothing with various methods and parameters, different normalisation methods and baseline corrections techniques. Binning involves modifying the resolution of the spectral data by replacing absorbance values of $n$ neighbouring wavelength by a single value, their mean; $n$ is referred to as the "binning factor". The binning operation is carried out by stepping through the dataset from the low to high wavelength in steps of $n$ such that each of the original wavelength only contributes to one bin. Smoothing was carried out using Savitzky-Golay filtering, wavelet denoising, and local polynomial fitting with gaussian weighting. The baseline correction methods included taking first and second derivatives of the spectra, a rubberband baseline correction, and a polynomial baseline correction.

Accordingly, models were created with different pre-processing methods in the sequence of normalisation, baseline correction, smoothing and binning. As PRFFECTv2 has been investigated previously with a similar dataset, the models were created based upon the best five pre-processing techniques combinations mentioned in the reference [22]. Table 1 shows the combinations of pre-processing techniques used with the SCIO and NIRscan Nano data sets in the binary and
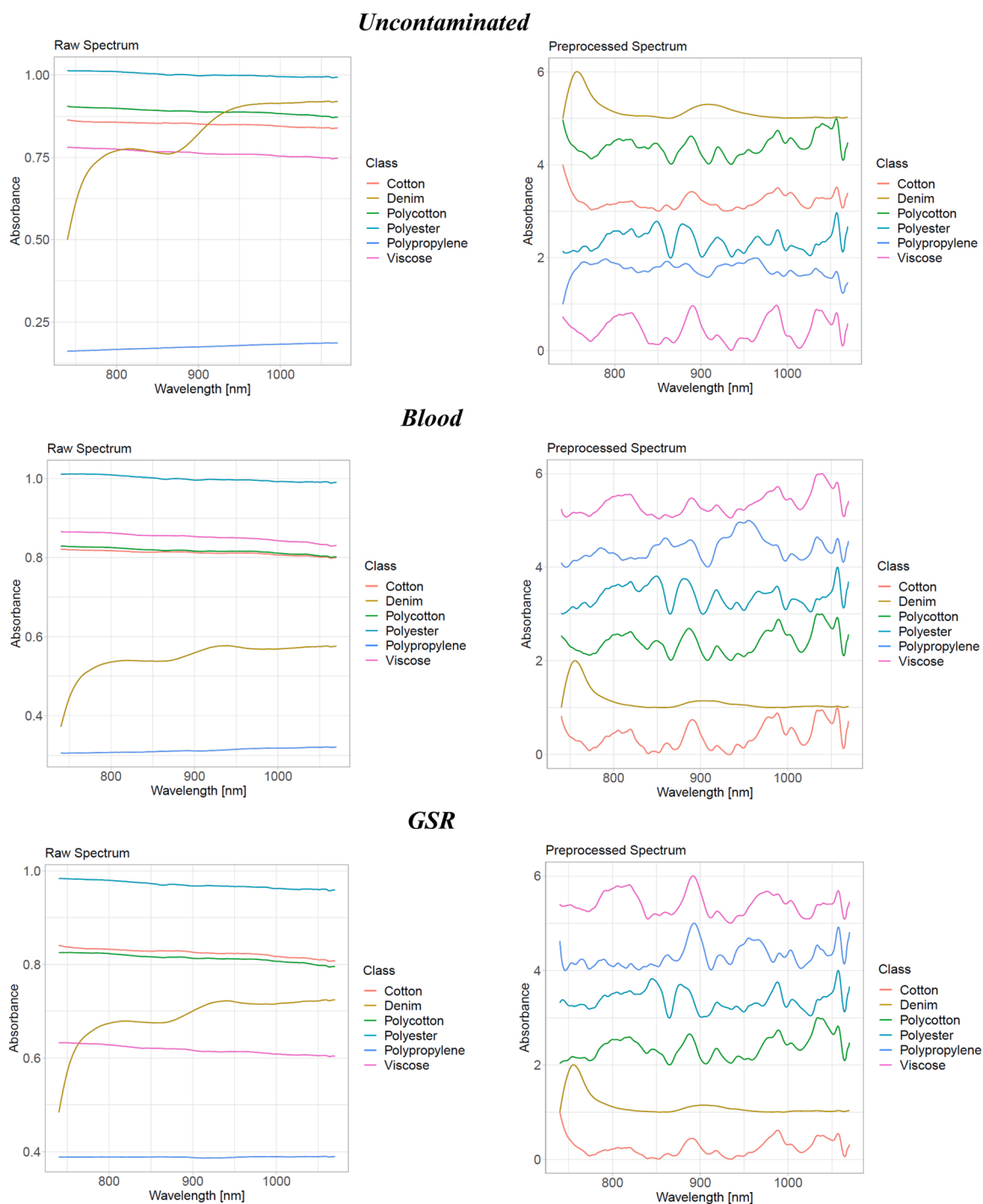
**Fig. 2.** The spectra of the six fabric materials after applying different contamination types obtained using the SCIO sensor. Left column is the raw spectra, and right column is the pre-processed with offset spectra.

multiclass-classification models.

## 3. Results and discussion

The raw spectra obtained from the SCIO spectrometer were pre-processed to reduce random variance caused by light scattering and enhance signal differences between different fabric types. Fig. 2 shows the raw and pre-processed (using baseline correction and smoothing) spectra of the fabric classes with the different contaminant types obtained using the SCIO spectrometer. The pre-processed spectra differ in

appearance from the raw spectra because of the use of a second derivative to reduce baseline effects. Spectra acquired in the 780 – 1060 nm region using the SCIO instrument contain peaks arising from the third overtone CH and second overtone OH vibrations.

Similarly, raw spectra acquired using the NIRscan Nano sensor were pre-processed using baseline correction and smoothing as shown in Fig. 3. The NIRscan Nano sensor measures spectra in the 900 – 1700 nm wavelength range, which contains information on the first and second overtones of molecular vibrations.

The clean fabrics NIR spectra was clearly shown peaks for the fabrics

## Sand
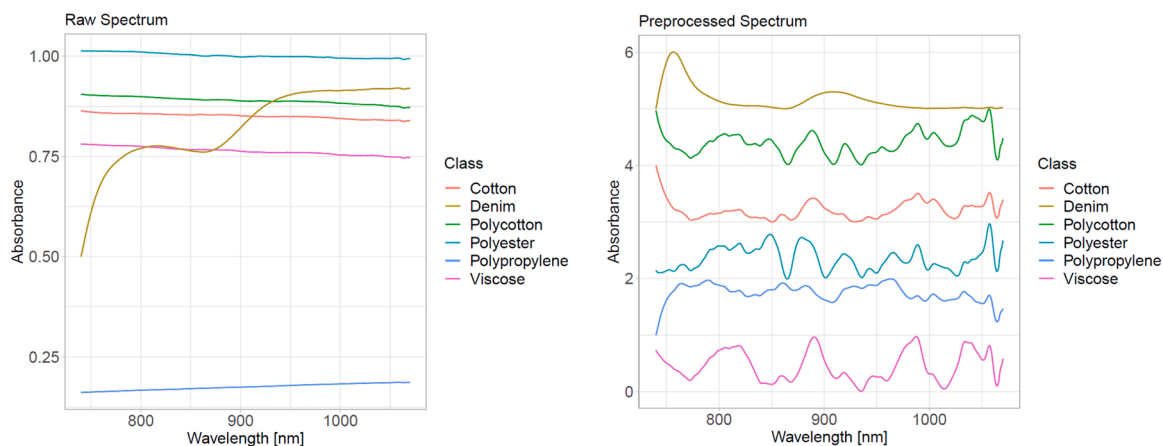


## Rainwater
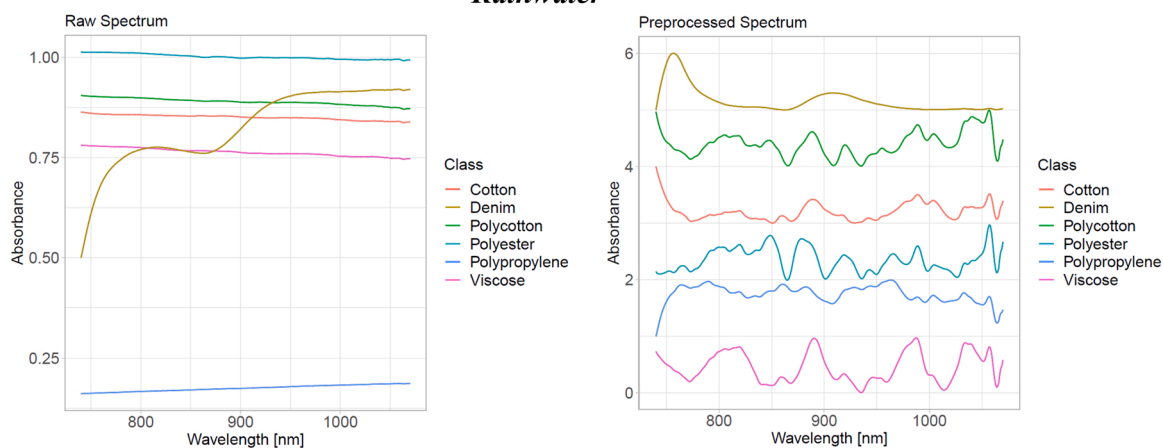


## Seawater
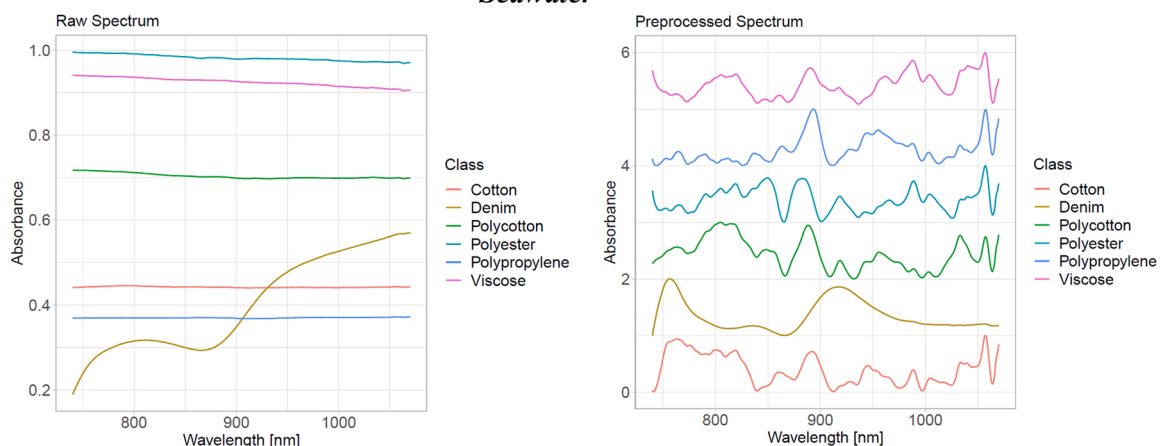


**Fig. 2.** (*continued*).

type, as for cellulose base fabrics (cotton, denim and viscose) peaks found at 1100 – 1300 nm (second overtone of CH stretching modes), 1350 – 1400 nm (combination band of the first overtone of CH stretching and the CH deformation vibrations) and 1400 – 1675 nm (first overtone of OH stretching modes in cotton and moisture) [19, 28-34]. The NIR spectra of polyester and the polycotton blend, which comprises 70% polyester, are comparable consisting of peaks at 1100 – 1250 nm (second overtone of CH stretching modes), 1350 – 1450 nm (combination band of the first overtone of CH stretching and the CH deformation vibrations, and first overtone of OH stretching vibration in

moisture) and the start of a peak at 1600 – 1750 nm (first overtone of CH stretching modes) [31,33,34]. The spectrum of polypropylene exhibits a broad peak at 1100 – 1300 nm (second overtone of CH stretching modes) [35]. The peak at approximately 925 nm in the spectra of all fabrics arises from third overtones of $CH_2$ stretching vibrations.

The NIR peaks of the contaminated fabrics exhibits similar peaks to the clean samples, which means it was hard to discriminate contaminant substances from the fabric's material. The NIR region for both spectrometers considered as short wavelength, this means that the light penetrates on fabrics increased which cause a weak and hard to

## *Uncontaminated*
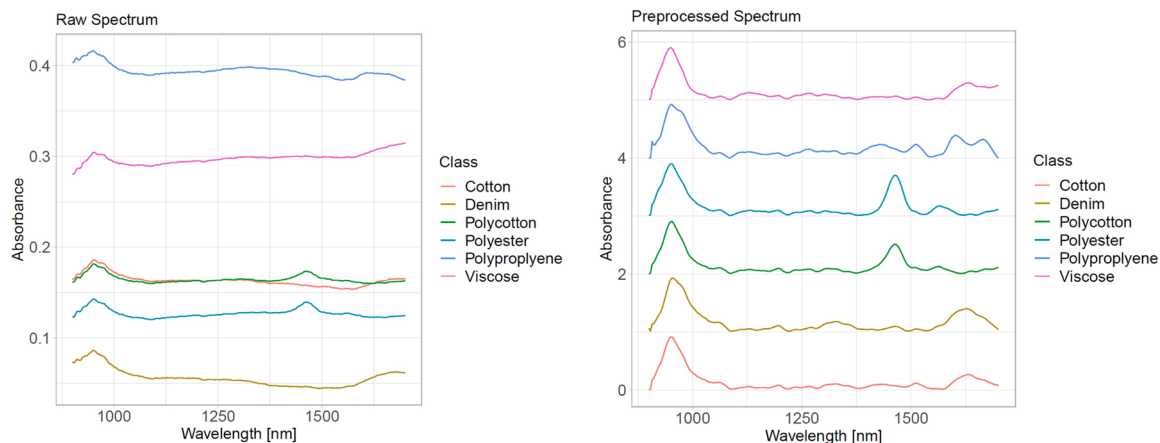


## *Blood*
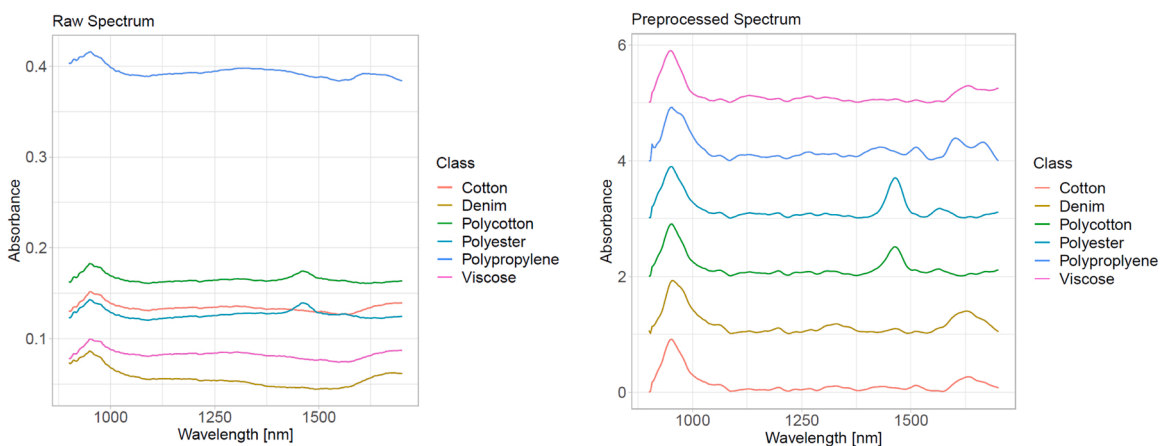


## *GSR*
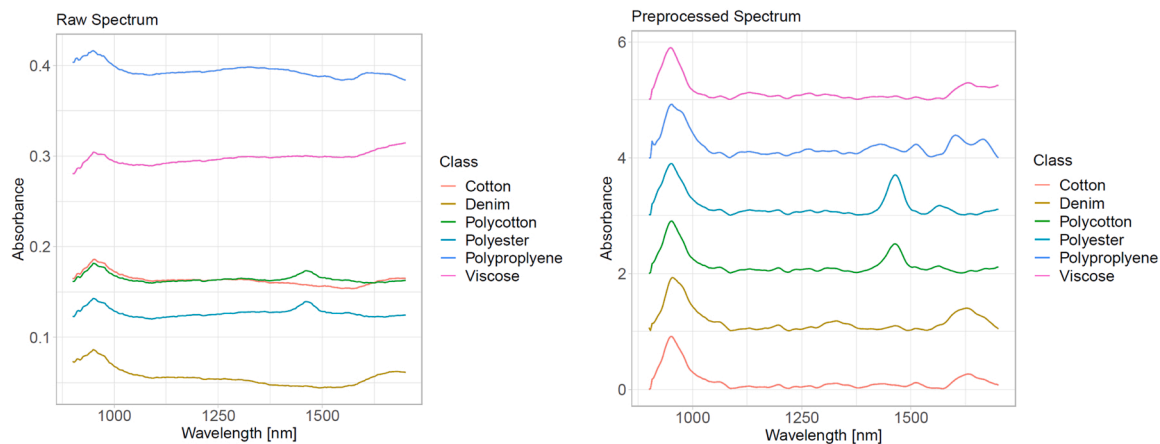


**Fig. 3.** The spectra of the six fabric materials after applying different contamination types obtained using the NIRnano scan spectrometer. Left column is the raw spectra, and right column is the pre-processed with offset spectra.

discriminate other contamination specially blood. Since only small amounts (50 μL) of blood were applied to the fabrics, protein or haemoglobin peaks were hard to distinguish in the spectra from both spectrometers (Fig. 2) [36]. The rainwater was rapidly absorbed by the fabrics, but subsequent drying meant that no significant OH peaks were detected. The seawater was absorbed quickly, but here there was a weak

signal which differed from uncontaminated spectra specially for polyester between 1000 and 1070 nm. Pure sand (Silicon-oxide) and GSR were also applied to the fabrics but did not produce any visible effects on the spectra. Finally, the water absorption on fabrics depends on the material type as its clearly shown by eye in the fabric samples used in this investigation.

## Sand
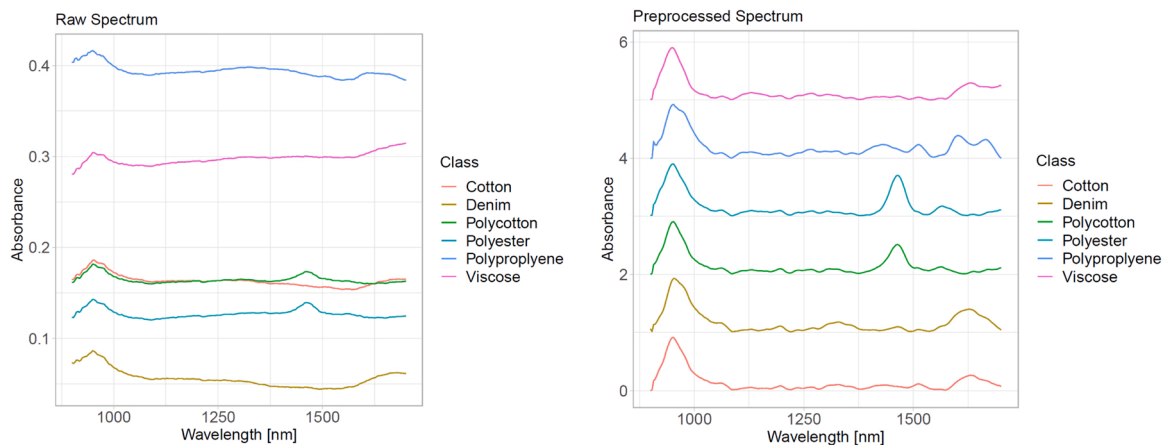


## Rainwater
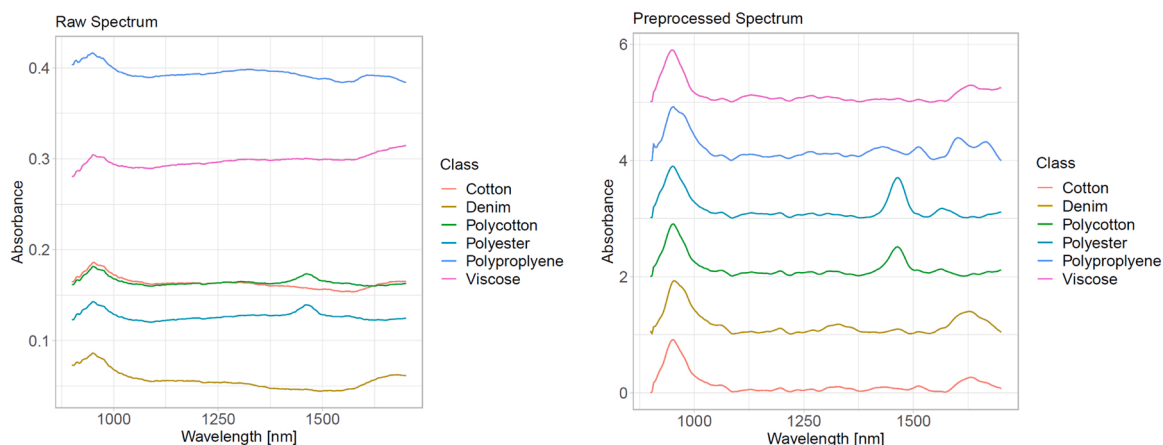


## Seawater
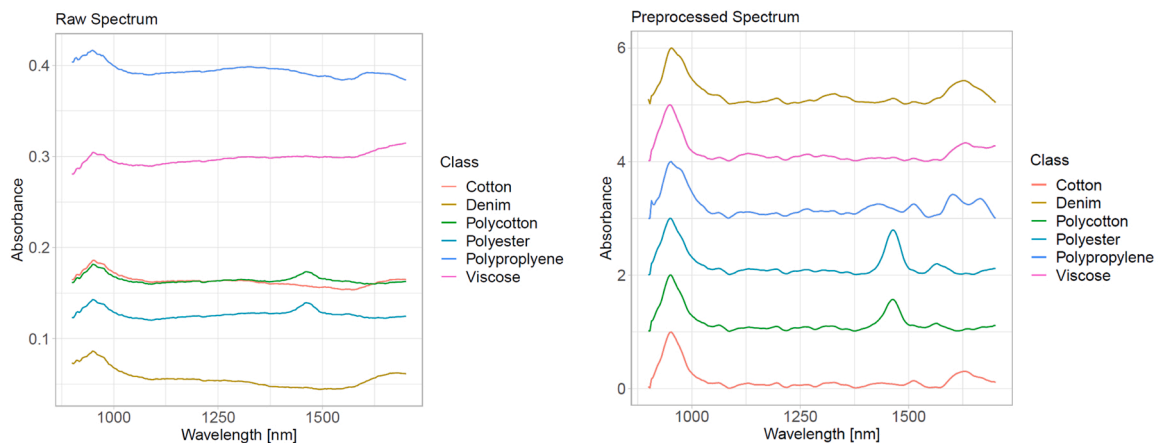


**Fig. 3.** (*continued*).

### 3.1. Data analysis and model validation

The random forest classifier in the PRFFECT v2 chemometric software was used to build binary and multiclass classification models for the SCIO and NIRscan Nano datasets to differentiate between fabrics types. The models were validated internally using 5-fold cross-validation on the training data, before predictions of a test set were made. The whole process was repeated 51 times on different training and test set partitions in order to assess the robustness of the models. The best five pre-processing combinations, used in a previous study on the identification of uncontaminated fabrics, were used to build the binary

and multiclass classification models [22]. The software generates feature importance plots to show the regions of the spectra that contribute to class differentiation as shown in Figs. 4 and 5 [25].

### 3.1.1. Binary classification

The five combinations of pre-processing methods used for the binary classification model are shown in Table 1. Tables 2 and 3 present the model statistics for classification "by spectra" and "by sample". The former are computed by treating each spectrum as an independent datapoint (although stratified sampling is employed to ensure that all spectra of the same sample are in the same training or testing set split, or
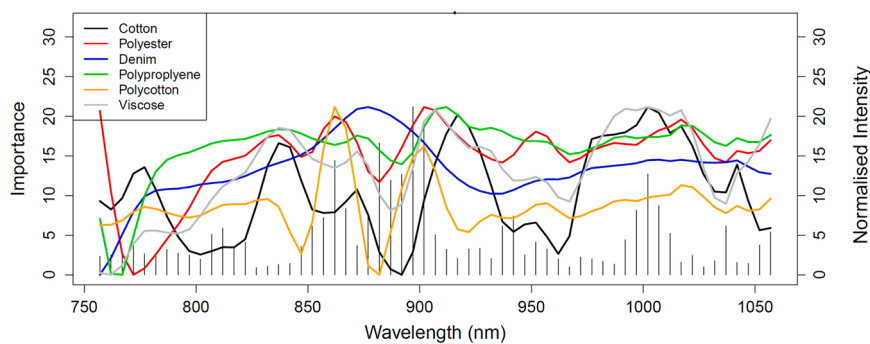
**Fig. 4.** The SCIO pre-processed spectra with importance features as Gini impurity of the best multiclass-classification model "2".
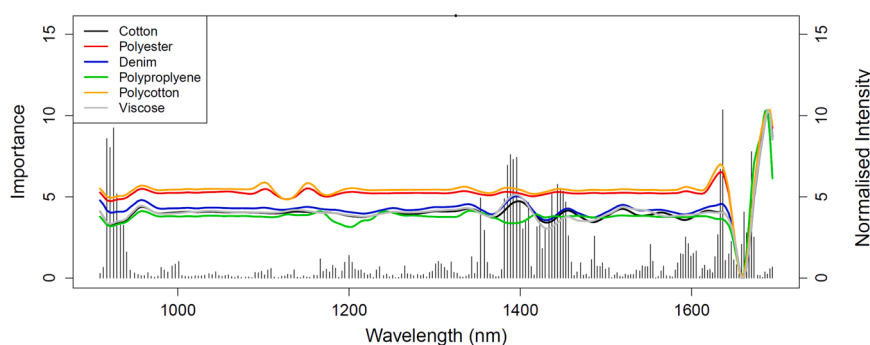


**Fig. 5.** The NIRScan Nano pre-processed spectra with importance features as Gini impurity of the best multiclass-classification model "1".

the same CV fold). The latter are calculated on a "by sample" basis by first taking the consensus (maximum vote) of the predictions of all spectra (five spectra per sample) for a given sample.

As shown in Table 2, the range of classification accuracies obtained using the SCIO method was 74–100%. The best model, Model 2, used pre-processing of binning by factor 5, vector normalisation, second derivative baseline correction and local polynomial fit with gaussian width 2 as the smoothing option. Some fabrics are similar in chemical composition, which makes it hard to differentiate between them. Denim is mainly cotton but differs in terms of the dye applied and viscose is a cotton-like structure material made by regenerated cellulose. Therefore, the classification statistics for identification of cotton, denim and viscose are lower than some other materials, but sensitivies of above 80% for both training and testing sets are still obtained [37,38]. Polyester, polycotton and polypropylene achieve an average classification accuracy of 95%, 89% and 86%, respectively.

The random forest classification models shown in Table 3 for the NIRscan Nano data achieved a range of accuracies from 86% to 100%. Model 1 was the optimum classification model, for which pre-processing was binning by factor 2, vector normalisation, and second derivative baseline correction and no smoothing option. As for the SCIO classification model, classifying denim from viscose and cotton was found to be more challenging than discriminating other fabrics. Polycotton was also observed to be misclassified as polyester in some instances. However, sensitivies of over 85% for cotton, viscose, polyester and polypropylene achieved an average classification accuracy of 88%, 90%, 95% and 98%, respectively [23,39].

The differences in classifying material class with each instrument was due to the differences in the wavelength ranges of the two spectrometers.

### 3.1.2. Multiclass classification

Multiclass classification models were trained to distinguish between cotton, denim, polyester, polycotton, polypropylene and viscose fabric materials. The mean classification results obtained for the pre-processed training and test sets with the SCIO and NIRScan instruments are shown in Tables 4 and 5, respectively. In general, the results obtained for a multi-class prediction model are clearly lower than those obtained using a binary classification model.

The optimum pre-processing combination for the SCIO and NIRscan Nano data were those employed in model 2 and model 1, respectively, which was binning with resolution reduction by five factors, local polynomial fit with Gaussian width 2 as smoothing choice, vector normalisation, and second derivative for baseline correction. The NIRscan Nano results show higher specificity than sensitivity for several of the classes. The only class to record a balanced sensitivity and specificity was polypropylene for which classification accuracies in the range of 98–100% were obtained. The statistics for by-spectra and by-sample classification modelling show exceptional performance. Furthermore, the results obtained using the SCIO instrument demonstrate higher specificity than sensitivity. The average accuracies obtained was 70–99%, with the lowest classification accuracy obtained for cotton 65–93%, and surprisingly, the highest classification accuracy was for polycotton 88–100%.

Overall, as expected the binary and multiclass class models show lower performance than the previous study due to the contamination but still it is in the very good range [22]. The SCIO spectrometer interrogates only the third overtone region, which provides limited information compared to the NIRscan Nano spectrometer, which interrogates the first and second overtone regions. Generally, the SCIO and NIRscan Nano spectrometers coupled with the spectral pre-processing and classification models builder tool (PRFFECTv2) enable rapid and non-destructive identification of contaminated textile fibre classes for

**Table 2**

The optimum binary classification models of the SCIO data. Accuracy, sensitivity and specificity are reported for 5-fold cross-validation (CV) and test set (TE) predictions calculated as means over 50 repeated experiments each using different stratified training and test set splits. The model numbers correspond to those given in Table 1.

| Model | Class | By Spectra | | | | | | By Sample | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Acc CV | Sens CV | Spec CV | Acc TE | Sens TE | Spec TE | Acc CV | Sens CV | Spec CV | Acc TE | Sens TE | Spec TE |
| 1 | Cotton | 0.837 | 0.868 | 0.831 | 0.807 | 0.79 | 0.81 | 0.871 | 0.925 | 0.859 | 0.83 | 0.828 | 0.831 |
| | PE | 0.865 | 0.855 | 0.868 | 0.849 | 0.835 | 0.853 | 0.885 | 0.89 | 0.88 | 0.871 | 0.887 | 0.868 |
| | Denim | 0.941 | 0.826 | 0.964 | 0.936 | 0.759 | 0.964 | 0.949 | 0.832 | 0.972 | 0.942 | 0.758 | 0.972 |
| | PP | 0.917 | 0.912 | 0.918 | 0.911 | 0.841 | 0.92 | 0.936 | 0.947 | 0.934 | 0.915 | 0.849 | 0.925 |
| | PC | 0.92 | 0.928 | 0.918 | 0.917 | 0.915 | 0.917 | 0.924 | 0.941 | 0.921 | 0.923 | 0.947 | 0.919 |
| | Viscose | 0.906 | 0.92 | 0.904 | 0.902 | 0.869 | 0.907 | 0.922 | 0.935 | 0.92 | 0.918 | 0.875 | 0.924 |
| 2 | Cotton | **0.904** | **0.928** | **0.9** | **0.884** | **0.87** | **0.886** | **0.92** | **0.955** | **0.914** | **0.899** | **0.875** | **0.903** |
| | PE | **0.842** | **0.863** | **0.838** | **0.801** | **0.788** | **0.804** | **0.874** | **0.919** | **0.864** | **0.818** | **0.819** | **0.818** |
| | Denim | **0.859** | **0.864** | **0.858** | **0.861** | **0.851** | **0.864** | **0.876** | **0.913** | **0.865** | **0.883** | **0.906** | **0.877** |
| | PP | **0.946** | **0.834** | **0.969** | **0.936** | **0.699** | **0.974** | **0.957** | **0.854** | **0.978** | **0.945** | **0.693** | **0.985** |
| | PC | **0.905** | **0.917** | **0.903** | **0.886** | **0.884** | **0.886** | **0.924** | **0.963** | **0.916** | **0.896** | **0.889** | **0.896** |
| | Viscose | **0.919** | **0.931** | **0.917** | **0.927** | **0.856** | **0.938** | **0.924** | **0.943** | **0.919** | **0.929** | **0.869** | **0.939** |
| 3 | Cotton | 0.817 | 0.857 | 0.807 | 0.802 | 0.767 | 0.809 | 0.839 | 0.888 | 0.828 | 0.819 | 0.779 | 0.829 |
| | PE | 0.845 | 0.859 | 0.842 | 0.843 | 0.851 | 0.84 | 0.858 | 0.9 | 0.846 | 0.857 | 0.917 | 0.844 |
| | Denim | 0.937 | 0.823 | 0.959 | 0.928 | 0.773 | 0.952 | 0.946 | 0.841 | 0.968 | 0.936 | 0.765 | 0.964 |
| | PP | 0.905 | 0.909 | 0.904 | 0.89 | 0.848 | 0.896 | 0.923 | 0.941 | 0.919 | 0.906 | 0.869 | 0.912 |
| | PC | 0.925 | 0.933 | 0.923 | 0.897 | 0.874 | 0.9 | 0.929 | 0.945 | 0.927 | 0.897 | 0.895 | 0.897 |
| | Viscose | 0.914 | 0.937 | 0.911 | 0.915 | 0.839 | 0.926 | 0.941 | 0.96 | 0.939 | 0.938 | 0.876 | 0.948 |
| 4 | Cotton | 0.826 | 0.842 | 0.823 | 0.786 | 0.758 | 0.792 | 0.854 | 0.873 | 0.851 | 0.801 | 0.765 | 0.809 |
| | PE | 0.854 | 0.878 | 0.847 | 0.848 | 0.817 | 0.856 | 0.87 | 0.919 | 0.856 | 0.861 | 0.843 | 0.866 |
| | Denim | 0.937 | 0.82 | 0.96 | 0.937 | 0.77 | 0.96 | 0.947 | 0.839 | 0.969 | 0.945 | 0.77 | 0.973 |
| | PP | 0.907 | 0.922 | 0.904 | 0.9 | 0.86 | 0.906 | 0.926 | 0.961 | 0.919 | 0.914 | 0.895 | 0.917 |
| | PC | 0.919 | 0.929 | 0.918 | 0.917 | 0.855 | 0.927 | 0.921 | 0.943 | 0.917 | 0.918 | 0.869 | 0.926 |
| | Viscose | 0.911 | 0.926 | 0.909 | 0.904 | 0.846 | 0.912 | 0.937 | 0.96 | 0.934 | 0.927 | 0.889 | 0.934 |
| 5 | Cotton | 0.786 | 0.836 | 0.775 | 0.75 | 0.759 | 0.747 | 0.798 | 0.849 | 0.787 | 0.756 | 0.745 | 0.759 |
| | PE | 0.892 | 0.877 | 0.896 | 0.891 | 0.863 | 0.897 | 0.917 | 0.934 | 0.911 | 0.911 | 0.906 | 0.913 |
| | Denim | 0.945 | 0.858 | 0.962 | 0.942 | 0.786 | 0.967 | 0.953 | 0.871 | 0.969 | 0.949 | 0.791 | 0.974 |
| | PP | 0.899 | 0.893 | 0.9 | 0.897 | 0.815 | 0.911 | 0.918 | 0.925 | 0.916 | 0.917 | 0.83 | 0.931 |
| | PC | 0.975 | 0.998 | 0.971 | 0.97 | 0.983 | 0.968 | 0.981 | 1 | 0.978 | 0.977 | 1 | 0.973 |
| | Viscose | 0.895 | 0.856 | 0.9 | 0.887 | 0.77 | 0.906 | 0.908 | 0.88 | 0.912 | 0.905 | 0.81 | 0.919 |

*In bold, the optimum model scores.

instant crime scene comparison.

## 4. Conclusion

This work shows the feasibility of using handheld Near IR spectrometers (SCIO® and NIRscan Nano) in conjunction with random forest models to classify contaminated fibre materials in a rapid, and non-destructive way for forensic comparison purposes at the crime scene.

The pocket-size spectrometers provide spectra in seconds and are non-destructive to the samples. Moreover, the random forest algorithm shows the potential of using NIR spectra from both spectrometers for textile fibre material classification. The contamination that was applied (blood, GSR, sand, rainwater and seawater) didn't affect the detection of the fabrics type on spectra visual comparison. The statistical program used was PRFFECTv2, which allowed a built-in algorithm starting from data pre-processing with numerous methods, flexibility in creating models, and is straight forward for real practical classification. The binary classification models built using NIR data acquired using the SCIO instrument show a prediction accuracy range of 88–99%, a sensitivity range of 78–100% and a specificity range of 87%– 100% for all fabric types. The multiclass-classification model has a prediction accuracy range of 76.9%– 86.7%, sensitivity range of 69.5%– 100% and specificity range of 92.9%– 100%. The binary classification models built using NIR data acquired using the NIRScan Nano instrument show a prediction accuracy range of 86–99%, a sensitivity range of 89 – 100% and a specificity range of 89–100%. The multiclass-classification model demonstrates a predication accuracy range of 82–86%, a sensitivity range of 73 – 100% and a specificity range of 94–100%. Overall, the scores were slightly less compared to previous work due to the contamination, but they still demonstrate the effectiveness of the methodology.

This study investigated the ability of the two micro-spectrometers to examine textile fibres and classification in-situ. However, further research is required to study those instruments in real-life crime scenes scenarios with different scenes, e.g. variations in weather and time of day, to investigate the possibility of taking good measurements with external light source. It is also recommended to investigate the timing of the contamination residues on the fibres which may have a different effect on the spectra.

## CRediT authorship contribution statement

**Hamad S. Rashed:** Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Andrew J. Parrott:** Methodology, Investigation, Writing – review & editing. **Alison Nordon:** Resources, Writing – review & editing, Supervision. **Matthew J. Baker:** Conceptualization, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition. **David S. Palmer:** Conceptualization, Software, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition..

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Table 3**

The optimum binary classification models of the NIRscan Nano data. Accuracy, sensitivity and specificity are reported for 5-fold cross-validation (CV) and test set (TE) predictions calculated as means over 50 repeated experiments each using different stratified training and test set splits. The model numbers correspond to those given in Table 1.

| Model | Class | By Spectra | | | | | | By Sample | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Acc CV | Sens CV | Spec CV | Acc TE | Sens TE | Spec TE | Acc CV | Sens CV | Spec CV | Acc TE | Sens TE | Spec TE |
| **1** | **Cotton** | **0.944** | **0.844** | **0.955** | **0.913** | **0.880** | **0.941** | **0.957** | **0.908** | **0.968** | **0.928** | **0.902** | **0.939** |
| | **PE** | **0.891** | **0.904** | **0.892** | **0.867** | **0.808** | **0.879** | **0.902** | **0.933** | **0.901** | **0.874** | **0.802** | **0.884** |
| | **Denim** | **0.894** | **0.879** | **0.891** | **0.912** | **0.932** | **0.915** | **0.901** | **0.899** | **0.902** | **0.932** | **0.927** | **0.939** |
| | **PP** | **0.999** | **1** | **0.992** | **1** | **1** | **1** | **0.997** | **1** | **0.997** | **1** | **1** | **1** |
| | **PC** | **0.881** | **0.831** | **0.9** | **0.877** | **0.788** | **0.893** | **0.896** | **0.879** | **0.893** | **0.89** | **0.845** | **0.891** |
| | **Viscose** | **0.965** | **0.978** | **0.943** | **0.933** | **0.892** | **0.954** | **0.832** | **0.994** | **0.954** | **0.944** | **0.898** | **0.954** |
| 2 | Cotton | 0.927 | 0.922 | 0.931 | 0.907 | 0.843 | 0.914 | 0.943 | 0.952 | 0.942 | 0.913 | 0.854 | 0.932 |
| | PE | 0.881 | 0.934 | 0.865 | 0.876 | 0.864 | 0.876 | 0.9 | 0.952 | 0.876 | 0.899 | 0.873 | 0.893 |
| | Denim | 0.934 | 0.923 | 0.911 | 0.9 | 0.849 | 0.916 | 0.922 | 0.935 | 0.915 | 0.904 | 0.832 | 0.922 |
| | PP | 0.99 | 0.998 | 0.991 | 0.991 | 0.965 | 0.992 | 0.998 | 1 | 0.998 | 0.992 | 0.973 | 0.995 |
| | PC | 0.867 | 0.801 | 0.887 | 0.864 | 0.801 | 0.866 | 0.89 | 0.854 | 0.891 | 0.866 | 0.833 | 0.862 |
| | Viscose | 0.954 | 0.99 | 0.959 | 0.951 | 0.934 | 0.943 | 0.965 | 0.99 | 0.954 | 0.943 | 0.942 | 0.944 |
| 3 | Cotton | 0.931 | 0.9 | 0.921 | 0.922 | 0.912 | 0.943 | 0.946 | 0.92 | 0.955 | 0.942 | 0.923 | 0.955 |
| | PE | 0.903 | 0.946 | 0.9 | 0.887 | 0.904 | 0.881 | 0.93 | 0.957 | 0.91 | 0.887 | 0.905 | 0.888 |
| | Denim | 0.925 | 0.922 | 0.916 | 0.917 | 0.876 | 0.927 | 0.93 | 0.938 | 0.925 | 0.919 | 0.855 | 0.93 |
| | PP | 0.99 | 1 | 0.991 | 0.991 | 0.982 | 1 | 0.999 | 1 | 0.999 | 0.999 | 0.986 | 1 |
| | PC | 0.9 | 0.833 | 0.911 | 0.901 | 0.785 | 0.921 | 0.919 | 0.858 | 0.927 | 0.911 | 0.79 | 0.925 |
| | Viscose | 0.961 | 0.992 | 0.955 | 0.952 | 0.943 | 0.951 | 0.954 | 1 | 0.959 | 0.952 | 0.926 | 0.953 |
| 4 | Cotton | 0.942 | 0.909 | 0.959 | 0.922 | 0.874 | 0.943 | 0.954 | 0.928 | 0.958 | 0.938 | 0.876 | 0.95 |
| | PE | 0.905 | 0.948 | 0.893 | 0.877 | 0.895 | 0.87 | 0.92 | 0.966 | 0.908 | 0.885 | 0.892 | 0.883 |
| | Denim | 0.931 | 0.906 | 0.935 | 0.923 | 0.819 | 0.94 | 0.936 | 0.911 | 0.94 | 0.922 | 0.803 | 0.942 |
| | PP | 0.991 | 0.99 | 0.99 | 0.987 | 0.961 | 0.991 | 0.991 | 1 | 0.991 | 0.992 | 0.965 | 0.987 |
| | PC | 0.9 | 0.822 | 0.901 | 0.845 | 0.8 | 0.899 | 0.905 | 0.855 | 0.904 | 0.9 | 0.826 | 0.908 |
| | Viscose | 0.954 | 0.982 | 0.958 | 0.96 | 0.865 | 0.955 | 0.954 | 0.989 | 0.96 | 0.961 | 0.878 | 0.977 |
| 5 | Cotton | 0.952 | 0.906 | 0.957 | 0.922 | 0.88 | 0.932 | 0.953 | 0.917 | 0.965 | 0.932 | 0.906 | 0.938 |
| | PE | 0.902 | 0.936 | 0.9 | 0.888 | 0.898 | 0.885 | 0.911 | 0.943 | 0.905 | 0.9 | 0.891 | 0.91 |
| | Denim | 0.915 | 0.893 | 0.931 | 0.928 | 0.809 | 0.953 | 0.931 | 0.897 | 0.939 | 0.936 | 0.803 | 0.955 |
| | PP | 0.992 | 0.992 | 0.992 | 0.99 | 0.983 | 0.999 | 0.993 | 1 | 0.993 | 0.991 | 0.976 | 1 |
| | PC | 0.91 | 0.826 | 0.924 | 0.9 | 0.781 | 0.915 | 0.921 | 0.845 | 0.924 | 0.9 | 0.776 | 0.913 |
| | Viscose | 0.954 | 0.988 | 0.958 | 0.959 | 0.954 | 0.957 | 0.965 | 1 | 0.958 | 0.961 | 0.946 | 0.958 |

*In bold, the optimum model scores.

**Table 4**

The optimum SCIO data multiclass classification results output, the mean accuracy, sensitivity and specificity scores across six different fabric materials with different pre-processing data methods used for random forest model method. The model numbers correspond to those given in Table 1.

| Model | Class | By Spectra | | | | | | By Sample | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Acc CV | Sens CV | Spec CV | Acc TE | Sens TE | Spec TE | Acc CV | Sens CV | Spec CV | Acc TE | Sens TE | Spec TE |
| **1** | **Cotton** | 0.789 | 0.72 | 0.942 | 0.801 | 0.742 | 0.945 | 0.819 | 0.765 | 0.951 | 0.834 | 0.799 | 0.953 |
| | **PE** | | 0.769 | 0.912 | | 0.806 | 0.914 | | 0.802 | 0.921 | | 0.838 | 0.922 |
| | **Denim** | | 0.81 | 0.973 | | 0.816 | 0.973 | | 0.834 | 0.976 | | 0.823 | 0.979 |
| | **PP** | | 0.839 | 0.976 | | 0.835 | 0.979 | | 0.867 | 0.982 | | 0.895 | 0.986 |
| | **PC** | | 0.805 | 0.955 | | 0.752 | 0.965 | | 0.839 | 0.956 | | 0.778 | 0.964 |
| | **Viscose** | | 0.809 | 0.982 | | 0.872 | 0.98 | | 0.823 | 0.992 | | 0.882 | 0.992 |
| **2** | **Cotton** | **0.788** | **0.727** | **0.937** | **0.787** | **0.787** | **0.934** | **0.819** | **0.758** | **0.944** | **0.816** | **0.784** | **0.978** |
| | **PE** | | **0.768** | **0.912** | | **0.754** | **0.901** | | **0.813** | **0.922** | | **0.863** | **0.983** |
| | **Denim** | | **0.826** | **0.976** | | **0.779** | **0.976** | | **0.804** | **0.979** | | **0.764** | **0.97** |
| | **PP** | | **0.842** | **0.975** | | **0.818** | **0.975** | | **0.874** | **0.984** | | **0.856** | **0.992** |
| | **PC** | | **0.792** | **0.96** | | **0.753** | **0.969** | | **0.843** | **0.962** | | **0.941** | **0.97** |
| | **Viscose** | | **0.817** | **0.98** | | **0.831** | **0.984** | | **0.832** | **0.986** | | **0.909** | **0.992** |
| **3** | **Cotton** | 0.792 | 0.707 | 0.941 | 0.808 | 0.726 | 0.943 | 0.821 | 0.727 | 0.949 | 0.843 | 0.799 | 0.948 |
| | **PE** | | 0.781 | 0.914 | | 0.824 | 0.929 | | 0.82 | 0.93 | | 0.853 | 0.941 |
| | **Denim** | | 0.829 | 0.979 | | 0.804 | 0.964 | | 0.852 | 0.979 | | 0.816 | 0.968 |
| | **PP** | | 0.832 | 0.973 | | 0.873 | 0.976 | | 0.876 | 0.977 | | 0.915 | 0.987 |
| | **PC** | | 0.803 | 0.958 | | 0.803 | 0.972 | | 0.843 | 0.959 | | 0.816 | 0.97 |
| | **Viscose** | | 0.817 | 0.979 | | 0.844 | 0.981 | | 0.818 | 0.987 | | 0.869 | 0.993 |
| **4** | **Cotton** | 0.783 | 0.707 | 0.937 | 0.808 | 0.762 | 0.941 | 0.819 | 0.757 | 0.945 | 0.836 | 0.779 | 0.946 |
| | **PE** | | 0.771 | 0.912 | | 0.808 | 0.926 | | 0.81 | 0.928 | | 0.853 | 0.941 |
| | **Denim** | | 0.8 | 0.974 | | 0.853 | 0.972 | | 0.819 | 0.977 | | 0.882 | 0.97 |
| | **PP** | | 0.844 | 0.975 | | 0.863 | 0.968 | | 0.884 | 0.981 | | 0.915 | 0.978 |
| | **PC** | | 0.785 | 0.958 | | 0.779 | 0.977 | | 0.83 | 0.958 | | 0.797 | 0.976 |
| | **Viscose** | | 0.808 | 0.978 | | 0.793 | 0.981 | | 0.821 | 0.987 | | 0.804 | 0.989 |
| **5** | **Cotton** | 0.781 | 0.722 | 0.952 | 0.789 | 0.712 | 0.939 | 0.819 | 0.796 | 0.958 | 0.837 | 0.779 | 0.952 |
| | **PE** | | 0.774 | 0.901 | | 0.827 | 0.91 | | 0.808 | 0.908 | | 0.863 | 0.923 |
| | **Denim** | | 0.793 | 0.976 | | 0.782 | 0.972 | | 0.823 | 0.979 | | 0.817 | 0.973 |
| | **PP** | | 0.831 | 0.968 | | 0.873 | 0.972 | | 0.832 | 0.982 | | 0.895 | 0.987 |
| | **PC** | | 0.777 | 0.957 | | 0.779 | 0.969 | | 0.834 | 0.962 | | 0.849 | 0.972 |
| | **Viscose** | | 0.799 | 0.977 | | 0.772 | 0.981 | | 0.829 | 0.987 | | 0.83 | 0.993 |

*In bold, the optimum model scores.

**Table 5**
The optimum NIRscan Nano data multiclass classification results output, the mean accuracy, sensitivity and specificity scores across six different fabric materials with different pre-processing data methods used for random forest model method.

| Model | Class | By Spectra | | | | | | By Sample | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Acc CV | Sens CV | Spec CV | Acc TE | Sens TE | Spec TE | Acc CV | Sens CV | Spec CV | Acc TE | Sens TE | Spec TE |
| **1** | **Cotton** | **0.816** | **0.874** | **0.974** | **0.853** | **0.857** | **0.972** | **0.835** | **0.892** | **0.978** | **0.867** | **0.872** | **0.969** |
| | **PE** | | **0.734** | **0.933** | | **0.821** | **0.959** | | **0.748** | **0.948** | | **0.833** | **0.974** |
| | **Denim** | | **0.719** | **0.965** | | **0.803** | **0.961** | | **0.737** | **0.967** | | **0.810** | **0.962** |
| | **PP** | | **0.990** | **0.998** | | **1.000** | **0.998** | | **0.992** | **0.998** | | **1.000** | **0.999** |
| | **PC** | | **0.695** | **0.936** | | **0.755** | **0.961** | | **0.752** | **0.938** | | **0.804** | **0.963** |
| | **Viscose** | | **0.912** | **0.970** | | **0.909** | **0.971** | | **0.915** | **0.917** | | **0.912** | **0.973** |
| **2** | Cotton | 0.835 | 0.863 | 0.978 | 0.838 | 0.873 | 0.974 | 0.848 | 0.874 | 0.980 | 0.851 | 0.887 | 0.975 |
| | PE | | 0.794 | 0.939 | | 0.774 | 0.949 | | 0.805 | 0.953 | | 0.779 | 0.963 |
| | Denim | | 0.732 | 0.967 | | 0.762 | 0.962 | | 0.732 | 0.965 | | 0.764 | 0.963 |
| | PP | | 0.987 | 0.999 | | 0.997 | 0.999 | | 0.993 | 1.000 | | 1.000 | 1.000 |
| | PC | | 0.738 | 0.946 | | 0.733 | 0.943 | | 0.792 | 0.949 | | 0.784 | 0.945 |
| | Viscose | | 0.912 | 0.970 | | 0.933 | 0.975 | | 0.913 | 0.970 | | 0.931 | 0.975 |
| **3** | Cotton | 0.834 | 0.874 | 0.974 | 0.843 | 0.836 | 0.982 | 0.846 | 0.888 | 0.974 | 0.855 | 0.857 | 0.980 |
| | PE | | 0.795 | 0.941 | | 0.780 | 0.943 | | 0.805 | 0.952 | | 0.789 | 0.956 |
| | Denim | | 0.735 | 0.970 | | 0.827 | 0.949 | | 0.737 | 0.971 | | 0.830 | 0.948 |
| | PP | | 0.989 | 0.999 | | 0.998 | 0.998 | | 0.993 | 1.000 | | 1.000 | 0.998 |
| | PC | | 0.716 | 0.945 | | 0.746 | 0.953 | | 0.752 | 0.947 | | 0.784 | 0.957 |
| | Viscose | | 0.912 | 0.968 | | 0.911 | 0.982 | | 0.913 | 0.969 | | 0.911 | 0.983 |
| **4** | Cotton | 0.833 | 0.853 | 0.975 | 0.844 | 0.875 | 0.967 | 0.845 | 0.868 | 0.976 | 0.853 | 0.887 | 0.966 |
| | PE | | 0.791 | 0.945 | | 0.823 | 0.944 | | 0.792 | 0.956 | | 0.828 | 0.955 |
| | Denim | | 0.745 | 0.961 | | 0.789 | 0.962 | | 0.767 | 0.961 | | 0.790 | 0.961 |
| | PP | | 0.987 | 0.998 | | 0.991 | 0.999 | | 0.990 | 0.998 | | 0.993 | 1.000 |
| | PC | | 0.725 | 0.946 | | 0.692 | 0.960 | | 0.764 | 0.946 | | 0.725 | 0.961 |
| | Viscose | | 0.908 | 0.972 | | 0.911 | 0.977 | | 0.910 | 0.973 | | 0.911 | 0.979 |
| **5** | Cotton | 0.817 | 0.873 | 0.973 | 0.845 | 0.867 | 0.977 | 0.828 | 0.890 | 0.975 | 0.852 | 0.877 | 0.976 |
| | PE | | 0.727 | 0.942 | | 0.766 | 0.956 | | 0.725 | 0.956 | | 0.769 | 0.967 |
| | Denim | | 0.790 | 0.975 | | 0.861 | 0.972 | | 0.796 | 0.976 | | 0.862 | 0.970 |
| | PP | | 0.976 | 0.981 | | 0.991 | 0.983 | | 0.990 | 0.979 | | 1.000 | 0.982 |
| | PC | | 0.662 | 0.933 | | 0.712 | 0.944 | | 0.705 | 0.932 | | 0.738 | 0.946 |
| | Viscose | | 0.902 | 0.972 | | 0.915 | 0.980 | | 0.891 | 0.972 | | 0.902 | 0.980 |

*In bold, the optimum model scores.

## Data Availability

Data will be made available on request.

## References

[1] R. Palmer, Identification and Comparsion, in: M.M. Houck (Ed.), Materials Analysis in Forensic Science, Elsevier, UK, 2016, pp. 195–204.

[2] R. Palmer, The Forensic Examination of Fibers, in: N.N. Daéid (Ed.), Interpols Forensic Science Review, researchgate, Lyon, France, 2010, pp. 182–190.

[3] E. Mistek, et al., Toward Locard's exchange principle: recent developments in forensic trace evidence analysis, Anal. Chem. 91 (1) (2019) 637–654, https://doi.org/10.1021/acs.analchem.8b04704.

[4] K. De Wael, L. Lepot, Forensic Science, Applications of Raman Spectroscopy to Fiber Analysis, in: J.C. Lindon, G.E. Tranter, D.W. Koppenaal (Eds.), Encyclopedia of Spectroscopy and Spectrometry, Third ed.,, Academic Press, Oxford, 2017, pp. 712–719.

[5] K. De Wael, K. Van Dijck, F. Gason, Discrimination of reactively-dyed cotton fibres with thin layer chromatography and UV microspectrophotometry, Sci. Justice 55 (6) (2015) 422–430, https://doi.org/10.1016/j.scijus.2015.06.001.

[6] T.G. Schotman, et al., Application of dye analysis in forensic fibre and textile examination: case examples, Forensic Sci. Int. 278 (2017) 338–350, https://doi.org/10.1016/j.forsciint.2017.07.026.

[7] Y. Ozaki, C.W. Huck, K.B. Beć, Chapter 2 - Near-IR Spectroscopy and Its Applications, in: V.P. Gupta (Ed.), Molecular and Laser Spectroscopy, Elsevier, 2018, pp. 11–38.

[8] C. Pasquini, Near infrared spectroscopy: a mature analytical technique with new perspectives – a review, Anal. Chim. Acta 1026 (2018) 8–36, https://doi.org/10.1016/j.aca.2018.04.004.

[9] F.B. de Santana, A.M. de Souza, R.J. Poppi, Visible and near infrared spectroscopy coupled to random forest to quantify some soil quality parameters, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 191 (2018) 454–462, https://doi.org/10.1016/j.saa.2017.10.052.

[10] S. Lee, et al., Random forest as a potential multivariate method for near-infrared (NIR) spectroscopic analysis of complex mixture samples: gasoline and naphtha, Microchem. J. 110 (2013) 739–748, https://doi.org/10.1016/j.microc.2013.08.007.

[11] C.S. Silva, A. Braz, M.F. Pimentel, Vibrational spectroscopy and chemometrics in forensic chemistry: critical review, current trends and challenges, J. Braz. Chem. Soc. (2019) https://doi.org/10.21577/0103-5053.20190140.

[12] X. Jin, et al., Spectral characterization and discrimination of synthetic fibers with near-infrared hyperspectral imaging system, Appl. Opt. 56 (12) (2017) 3570–3576, https://doi.org/10.1364/AO.56.003570.

[13] W.D. Li, X.H. Wang, L.H. Peng, Identification of natural bamboo fibers and flax fibers, Adv. Mater. Res. 821–822 (2013) 153–157, https://doi.org/10.4028/www.scientific.net/AMR.821-822.153.

[14] R.M. Correia, et al., Banknote analysis by portable near infrared spectroscopy, Forensic Chem. 8 (2018) 57–63, https://doi.org/10.1016/j.forc.2018.02.003.

[15] Guifang, W., M. Hai, P. Xin. Identification of varieties of natural textile fiber based on Vis/NIR spectroscopy technology. in 2015 IEEE Advanced Information Technology, Electronic and Automation Control Conference (IAEAC). 2015. DOI: ⟨https://doi.org/10.1109/IAEAC.2015.7428621⟩.

[16] H. Chen, Z. Lin, C. Tan, Classification of different animal fibers by near infrared spectroscopy and chemometric models, Microchem. J. 144 (2019) 489–494, https://doi.org/10.1016/j.microc.2018.10.011.

[17] H. Chen, C. Tan, Z. Lin, The feasibility study of non-destructive detection of cashmere by near-infrared spectroscopy and data driven-based class-modeling, Vib. Spectrosc. 102 (2019) 57–62, https://doi.org/10.1016/j.vibspec.2019.04.006.

[18] J. Zhou, et al., Fiber-content measurement of wool–cashmere blends using near-infrared spectroscopy, Appl. Spectrosc. 71 (10) (2017) 2367–2376, https://doi.org/10.1177/0003702817713480.

[19] Y. Liu, et al., Variation and relationship of quality and near infrared spectral characteristics of cotton fibers collected from multi-location field performance trials, Text. Res. J. 85 (14) (2015) 1474–1485, https://doi.org/10.1177/0040517514566106.

[20] W. Jiang, et al., Classification and identification of plant fibrous material with different species using near infrared technique—a new way to approach determining biomass properties accurately within different species, Front. Plant Sci. 7 (2017), https://doi.org/10.3389/fpls.2016.02000.

[21] A.V. Morillas, J. Gooch, N. Frascione, Feasibility of a handheld near infrared device for the qualitative analysis of bloodstains, Talanta 184 (2018) 1–6, https://doi.org/10.1016/j.talanta.2018.02.110.

[22] H.S. Rashed, et al., A comparative investigation of two handheld near-ir spectrometers for direct forensic examination of fibres in-situ, Vib. Spectrosc. 113 (2021), 103205, https://doi.org/10.1016/j.vibspec.2020.103205.

[23] Baratloo, A., et al., Part 1: simple definition and calculation of accuracy, sensitivity and specificity. 2015. ⟨https://doi.org/10.22037/emergency.v3i2.8154⟩.

[24] Z. Xiaobo, et al., Variables selection methods in near-infrared spectroscopy, Anal. Chim. Acta 667 (1) (2010) 14–32, https://doi.org/10.1016/j.aca.2010.03.048.

[25] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32, https://doi.org/10.1023/A:1010933404324.

[26] D.S. Palmer, et al., Random Forest Models to Predict Aqueous Solubility, J. Chem. Inf. Model 47 (1) (2007) 150–158, https://doi.org/10.1021/ci060164k.

[27] Smith, B.R., M.J. Baker, D.S. Palmer, PRFFECT: a versatile tool for spectroscopists. 2018. ⟨https://doi.org/10.1016/j.chemolab.2017.10.024⟩.

[28] J. Rodgers, K. Beck, NIR characterization and measurement of the cotton content of dyed blend fabrics, Text. Res. J. 79 (8) (2009) 675–686, https://doi.org/10.1177/0040517508090884.

[29] J. Rodgers, et al., Near Infrared Measurment of Cotton Fiber Micronaire, Maturity and Fineness - a Comparative Investigation, Text. Res. J. 80 (2010) 780–793, https://doi.org/10.1177/0040517509343780.

[30] J. Rodgers, J. Zumba, C. Fortier, Measurement comparison of cotton fiber micronaire and its components by portable near infrared spectroscopy instruments, Text. Res. J. 87 (1) (2017) 57–69, https://doi.org/10.1177/0040517515622153.

[31] C. Ruckebusch, et al., Quantitative analysis of cotton—polyester textile blends from near-infrared spectra, Appl. Spectrosc. 60 (5) (2006) 539–544, https://doi.org/10.1366/000370206777412194.

[32] J. Zumba, J. Rodgers, Cotton micronaire measurements using small portable near-infrared (NIR) analyzers, Appl. Spectrosc. 70 (5) (2016) 794–803, https://doi.org/10.1177/0003702816638227.

[33] Y. Liu, et al., Least-squares support vector machine and successive projection algorithm for quantitative analysis of cotton-polyester textile by near infrared spectroscopy, J. Infrared Spectrosc. 26 (1) (2018) 34–43, https://doi.org/10.1177/0967033518757069.

[34] J. Workman Jr, L. Weyer, Practical Guide and Spectral Atlas for Interpretive Near-infrared Spectroscopy, CRC Press, 2012.

[35] Zhou, J., et al., Textile Fiber Identification Using Near-Infrared Spectroscopy and Pattern Recognition. 2018.

[36] S.A. DeJong, et al., Detection limits for blood on four fabric types using infrared diffuse reflection spectroscopy in mid- and near-infrared spectral windows, Anal. Chem. 87 (17) (2015) 8740–8747, https://doi.org/10.1021/acs.analchem.5b01825.

[37] R. Paul, Denim: Manufacture, Finishing and Applications, Elsevier, 2015.

[38] J. Chen, Chapter 4 - Synthetic Textile Fibers: Regenerated Cellulose Fibers, in: R. Sinclair (Ed.), Textiles and Fashion, Woodhead Publishing, 2015, pp. 79–95.

[39] J. Robertson, C. Roux, Fibers: overview, in: M.M. Houck (Ed.), Materials Analysis in forensic science, Elsevier, UK, 2016, pp. 171–175.