Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Soil provenance using ATR-FTIR spectroscopy augmented with chemometrics–A pilot study

Praveen Kumar Yadav
National Forensic Sciences University
Sweety Sharma
Sharmasweety511@gmail.com

Manav Rachna International Institute of Research and Studies https://orcid.org/0000-0002-7886-4672

Research Article

Keywords: Forensic Science, ATR-FTIR, Soil, Chemometrics, PCA, PLS-DA, Geographical discrimination, Soil characterization

Posted Date: August 14th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10693333/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Title–Soil provenance using ATR-FTIR spectroscopy augmented with chemometrics–A pilot study

Abstract

ATR-FTIR spectroscopy coupled with chemometric tools was employed for the analysis. The spectral data obtained from ATR-FTIR were further processed using multivariate statistical techniques such as Principal Component Analysis (PCA) and Partial Least Squares Discriminant Analysis (PLS-DA) to interpret the results and classify the samples. The discrimination of soil samples was achieved based on variations in their organic and inorganic constituents. The PCA model achieved an overall classification accuracy of 86.11%, and the PLS-DA model showed an accuracy of 85.56%, indicating good discriminatory capability among the analysed classes. The present study effectively highlights the potential of ATR-FTIR spectroscopy combined with chemometric analysis as a reliable, rapid, and non-destructive approach for the geographical discrimination of soil samples.

Keywords: Forensic Science; ATR-FTIR; Soil ; Chemometrics; PCA; PLS-DA; Geographical discrimination; Soil characterization

1. Introduction

Forensic examination of soil has a vivid history discussed prominently in the writings of Sir Aurthur Conan Doyle between 1887 and 1927. Hans Gross, in 1893, was first to suggest that “perhaps the dirt on someone’s shoes could tell more about where a person had last been than toilsome inquirie”[1]. In 1908, Georg Popp, a German chemist, used forensic examination of soil to assist the criminal investigation of Margarethe Filbert murder. Since then, the science of soil analysis has come a long way[2]. Forensic examination initiated with the visual and microscopic examination has now advanced into the organic and inorganic analysis, qualitative and quantitative analysis, and comparison studies.

Pedology or the science of soil deals with different kinds of soil, their distribution, factors affecting its composition, sampling and description of soil, quality of soil samples, its spatial variability, and the interpretation of soil analysis[3]. Soil is a complex mixture of various organic and inorganic components which depends on a diverse array of elements, such as the geographical features of the terrain, the prevailing atmospheric conditions, the progression of time, the biological activities of plants and microorganisms, the characteristics of water, and indeed, the influence of human actions. Soil diversity might also vary depending upon the fossilization of soil along with the pollen and spores found at the site, making forensic soil examination an intricate task due to its inherent complexity. This variation in soil composition can be used to discriminate soil samples from different sites and in turn to accomplish the soil provenance[4].

Forensic examination of soil is a new method-oriented, technique driven activity often utilizing the principles of paedology, soil survey, soil mineralogy, soil chemistry, soil molecular biology, soil geophysics, and forensic science[5–7]. Primary and secondary transfer of soil traces, commonly found on shoes and clothing, are dictated by the Edmond's principle of mutual exchange[8]. Soil evidence can be important in criminal cases such as hit and run case, sexual assault case, abduction, murder, wildlife crime, theft, etc.; in civil cases such as drug shipment and construction; and intelligence cases to identify the primary and secondary crime scenes[9]. Organic components of soil have been previously characterized using gas chromatography – mass spectroscopy[10–12], liquid chromatography[13,14], ion chromatography[15,16], Raman spectroscopy[17–19], and infrared spectroscopy[20,21]. Inorganic components of soil are analysed using atomic absorption spectroscopy[22], Energy dispersive X-ray diffraction spectroscopy[23–25], Inductively coupled plasma – mass spectroscopy and inductively coupled plasma – atomic emission spectroscopy[26–28].

Forensic examination of soil accomplices three goals – a) linking crime scene, accused, and victim, b) geographic provenancing and source attribution, and c) crime scene reconstruction. A connection can be established between the soil traces found at the crime scene, on the victim, and belonging to the accused by undertaking a comparative analysis of their respective compositional profiles. Similarly, by searching and comparing the profiles of questioned and reference samples, geographic origin of the samples can be ascertained. By studying the circumstances of soil trace transfer, it can help in crime scene reconstruction[7,29].

In the present pilot study, an attempt has been made to discriminate soil samples collected from various sites using attenuated total reflectance – Fourier transform infrared (ATR-FTIR) spectroscopy augmented with chemometric tools. ATR-FTIR spectroscopy is based on the total internal reflection of the incident light. The interaction with evanescent wave produces a spectrum that represents molecular vibration characteristic of specific bonds. The sample is placed directly over the high refractive index crystal, made up of diamond, ZnSe, or germanium[30]. This shallow penetration depth is advantageous in forensic soil analysis as it minimizes bulk interference and allows surface-specific characterization[31]. ATR-FTIR spectroscopy has emerged as a non-destructive, rapid, less time consuming, and cheaper alternative to conventional techniques such as GC-MS, and HPLC. However, overlapping bands, surface biased analysis, and effect of environmental factor necessitate optimal sample preparation methods[32–34].

2. Materials and Methods

2.1 Sample Collection

Soil samples were collected from 15 different geographical regions of India, as detailed in Table 1. From each location, 6 representative samples were collected, resulting in a total of 90 samples.

Table 1: Soil samples collected from different states

Sample No.	Location	State	Latitude (N)	Longitude (E)
01	Bangoru Falls	Jammu & Kashmir	32.0526	76.7323
02	Junga-	Himachal Pradesh	31.0245	77.1916
03	Mohali-Shimla Road	Punjab/Himachal Pradesh	30.68	76.72
04	Mohan Garden, New Delhi	Delhi	28.6211	77.0405
05	Tara Devi	Himachal Pradesh	31.0539	77.1419
06	Bir Billing, Baijnath	Himachal Pradesh	32.0598	76.7445
07	Angamaly	Kerala	10.1926	76.3869
08	CUSAT (Cochin University Area)	Kerala	10.0531	76.3528
09	Ernakulam	Kerala	9.98	76.3
10	Kannur	Kerala	11.8705	75.3679
11	Pala, Kottayam	Kerala	9.5916	76.5222
12	Palakkad	Kerala	10.7733	76.6580
13	Pali Malli	Kerala	30.1105	79.3930
14	Perumbavoor	Kerala	10.1146	76.4778
15	Thrissur	Kerala	10.5241	76.2121

2.2 Sample Preparation

All Soil samples were collected from different regions and dried at room temperature. The debris, leaves, stones, and other contaminants were removed from the soil using the 2 mm standard sieve. After the sieving process, fine particles were collected and stored in zipper bags and analysed. In total 90 samples (6 from each site) were collected and analysed. The fine particles of soil samples were scanned using ATR FT-IR spectroscopy. The repeatability test was carried out by analysing a single soil sample three times. A reproducibility test was conducted by analysing 3 samples of soil collected at distinct points of one location.

2.3 Instrumentation

Samples were analysed by using the 'Nicolet iS20 ATR-FTIR spectrophotometer (Perkin Elmer) equipped with an iTX Smart ATR accessory and diamond crystal. The spectral range selected was 4000-600 cm^{-1} (mid-infrared range) over the resolution of 4 cm^{-1}. A total of 32 scans were taken and the final acquired spectrum was the average of the accumulated scans. To apply the homogenous pressure, samples were pressed maximum using the knob and anvil of ATR. Before analyzing any new sample, the ATR diamond crystal was cleaned thoroughly using spectroscopic-grade Acetone. All measurement scans were conducted at room temperature (25±5ºC).

2.3 Chemometrics

2.4.1 PCA

PCA is a dimensionality reduction tool for high-dimensional data sets such as of spectroscopic analysis and is consequently used to spot the samples' spectral spatial distribution. It is executed by decomposing or reducing the linear combination of variables into a few principal components. The obtained results with the initial 3 PCs (PC1, PC2, and PC3) were assessed. However, the first 2 PCs (PC1, and PC2) were considered the most representative of the analysed data [35].

2.4.2 **PLS-DA (Partial least squares discriminant analysis)**

PLS-DA was carried out to develop a model for the classification of soil samples. This method is entirely dependent on the partial least squares regression approach, used to correlate the dependent (Y) and independent variables (X). The X variables were the spectral matrix; however, the Y variables were the classes of selected timber samples. For the evaluation of PLSR regressions, the RMSE, that is, root mean square error, was calculated. The RMSE value quantifies the effectiveness of the calibration model to predict the property of interest in various samples of unknown class [36].

2.4.3 Spectral Pre-Processing

All the obtained spectra were pre-processed before going to the chemometric analysis. Different preprocessing methods were applied in order to enhance the information contained in the spectral data and to reduce the noise ratio. The following are the parameters and pre-processing methods used to build the PCA, PLS-DA, and LDA models. All spectral data were mean-centred before analysis. Pre-processing techniques such as baseline correction and smoothing (Savitzky–Golay) were applied in PCA to improve signal quality, whereas baseline correction and deresolution were applied in PLS-DA to enhance discriminatory power between sample classes. The detailed pre-processing methods and parameters are given in Table 2

Table 2: Pre-processing methods for PCA, PLS-DA and LDA

Parameter	PCA Model	PLS-DA Model
Pre-processing Applied	Baseline Correction; Savitzky–Golay Smoothing	Baseline Correction; Deresolve Transformation
Smoothing	Savitzky–Golay	Not Applied
De-resolution	Not Applied	Applied

Mean Centering	Applied	Applied
Algorithm Used	NIPALS	Kernel
Validation Method	Cross-validation	Cross-validation
Cross-validation Type	Random (20 segments)	Random (20 segments)
Total Components / Factors	7 Components	7 Factors

3. Results and discussion

In the present study, ATR FT-IR spectroscopy with chemometrics tools, was used for the differentiation of 90 soil samples collected from 15 different geographical regions (6 samples from each site). Figure 1 illustrates the representative spectra of soil and corresponding peaks are mentioned in Table 3. The results of reproducibility and repeatability showed an exact overlapping of spectra with each other without any significant variations. The standard deviation of ± 0.18 and ±0.19 was calculated for repeatability variation test and reproducibility test respectively of three full range spectral data points.

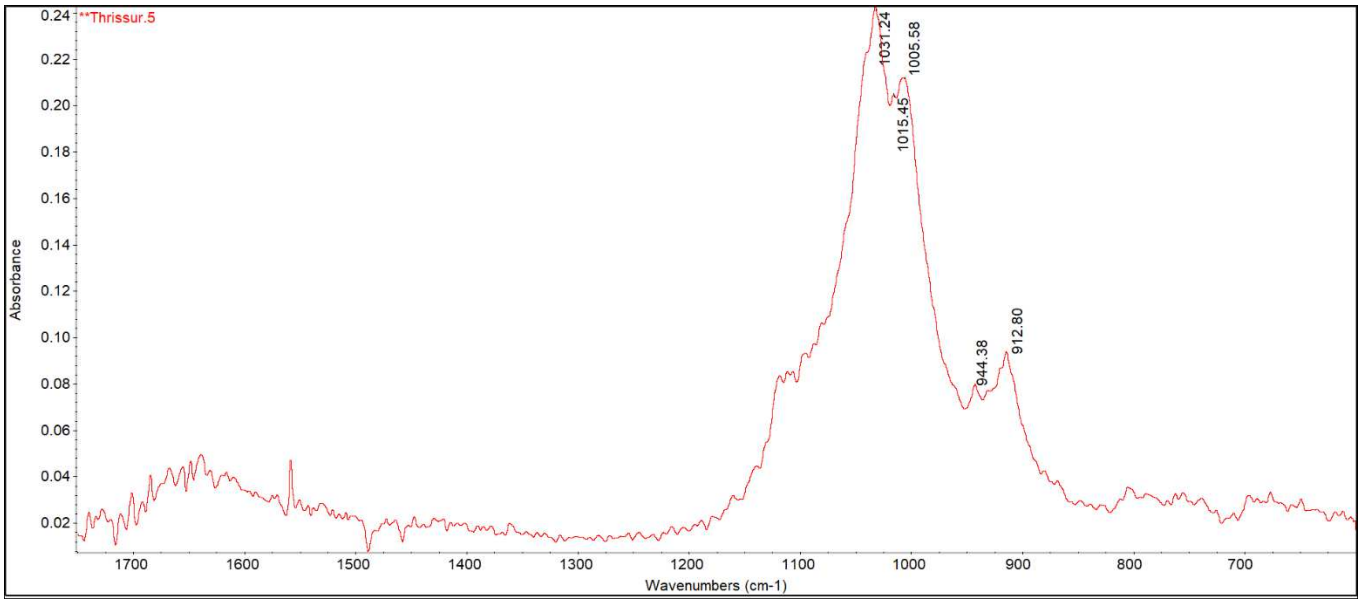


Figure 1 Representative spectrum of soil (Thrissur)

3.1 Visual Discrimination (1300-600cm^{-1}):

The major peaks in the spectra of the soil sample are observed in the range of 1010 - 1045 cm^{-1} and 870 - 954 cm^{-1}. Peak positioned at **1033 cm^{-1} (**1045-1010 cm^{-1}**) corresponds to** Al-OH deformation (kaolinite) with vibrational band of RHC=CH2, and aromatic C-H out-of-plane bending. Peak positioned in the range of 870- 945 cm^{-1} corresponds to Benzoic acid, pyranose ring (carbohydrates), cellulose. The details of other peaks are elaborated in Table 3.

Table 3: Vibrational peaks (Reported peaks in literature)

Peak Region (cm^{-1})	Inorganic Constituents	Organic Constituents
3625–3615	Si–O–H vibrations (clays, gibbsite, iron oxides, kaolinite), free O–H, N–H stretching	Moisture, oxygen-containing organic compounds
3440–3320	Hydrogen-bonded O–H and N–H stretching	—
2965–2853	—	Aliphatic C–H symmetric and asymmetric stretching
1660–1640	O- and N-containing polar functional groups	Amide C=O stretching, aromatic C=C stretching, carboxylate C–O asymmetric stretching, C=N stretching, conjugated ketone C=O
1444–1408	—	Aliphatic C–H bending

1403–1354	—	C–O stretching, O–H deformation (COOH), phenolic C–O stretching
1185–1144	—	C–OH of aliphatic alcohols, carbohydrates
1045–1010	Al–OH deformation (kaolinite)	Aromatic C–H out-of-plane bending, $RHC=CH_2$
1005–995	SiO_2, Si–O lattice stretching	Polysaccharides, aromatic =CH, C=C groups
945–870	—	Benzoic acid, pyranose ring (carbohydrates), cellulose, $RHC=CH_2$, $R_2C=CH_2$
870–890	Carbonates	Lignins (aromatic C–H out-of-plane)
820–752	Inorganic materials (clay, quartz, carbonates, kaolinite)	$R_2C=CHR$, phenyl group
725–720	Calcite	Long-chain alkanes (C–H rocking, methyl)
697–690	SiO_2	—
~650	Bentonite	—
535–525	Si–O–Al deformation (kaolinite)	—
470–460	Si–O–Si bending (illite, smectite, kaolinite)	—

3.2. Soil discrimination using PCA model

Principal component analysis is a data reduction technique which is used to study the relative positioning of the samples on the axis of principal components. Principal component

analysis consists of a score plot and a loading plot. The score plot demonstrates that Principal Component 1 (PC-1) accounts for 99% of the total variance, indicating that the majority of differences among the samples are governed by a single dominant factor. However, PC-2 (1%) contributes negligibly, suggesting minimal additional variation beyond that captured by PC-1. Fig. 1 shows the Explained variance plot of PCs, indicating that PC1 accounts for most of the variation, and beyond PC1, a flat line is observed.

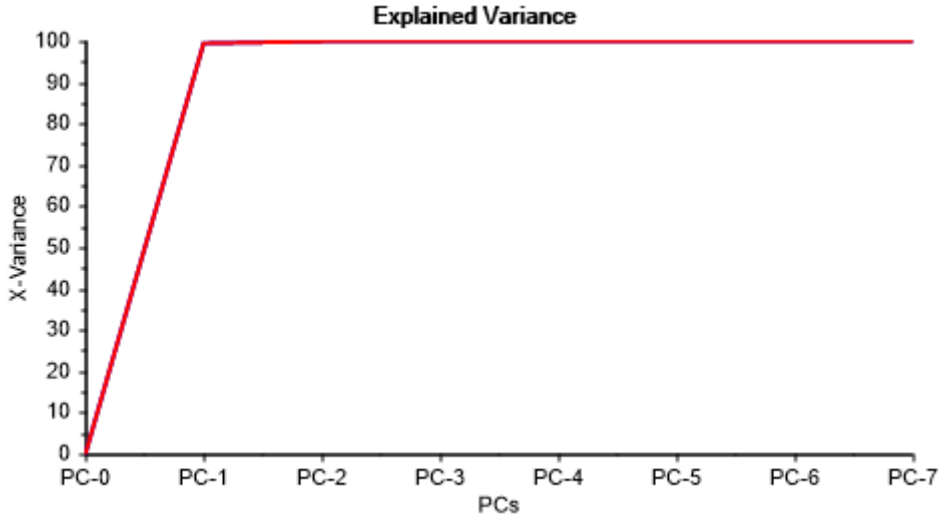


Fig. 3 Explained variance plot of PCs

PCA score plot, as shown in figure 4, illustrates distinct and well-defined sample clusters, with points within each group closely packed, reflecting high inter-group variation, high intra-group similarity, and good reproducibility. Out of 15 classes, 6 classes were easily separated out on the score plot (classes 1 to 6). However, classes 1 and 4 were closely placed to each other. The remaining classes, including the samples from sample numbers 7-15 (7-Angamaly; 08- CUSAT; 09-Ernakulam; 10- Kannur; 11- Pala, Kottayam; 12- Palakkad; 13-Pali Malli; 14- Perumbavoor; 15- Thrissur) were closely placed. This might be close provenance of these samples as well as the local geological formations, environmental conditions, or anthropogenic factors. Given that the collected samples were representative of

the common state of Keralam, a subsequent step involved their separation, followed by further discrimination utilizing the model that had been developed.

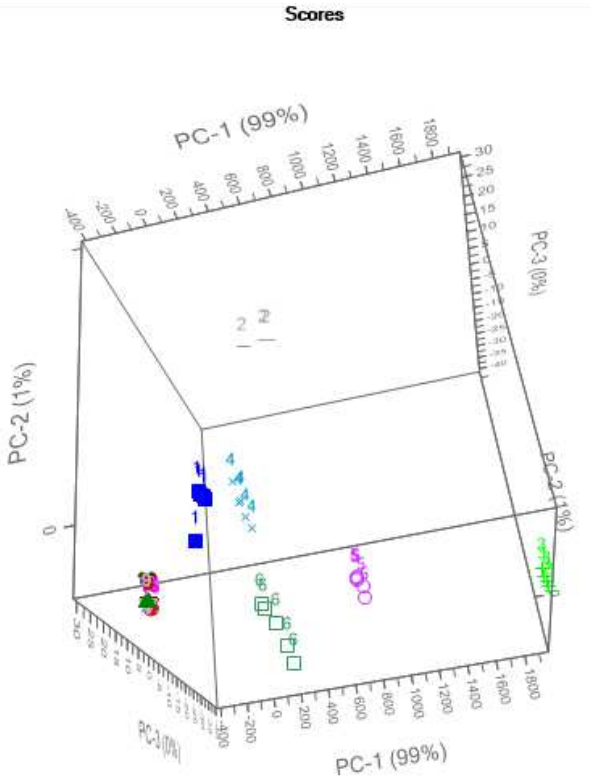


Fig. 4: 3-D PCA score plot to differentiate the soil samples collected from different geographical regions

The PCA loading plot (Fig 5) shows relatively consistent positive loadings across most of the spectral range, indicating that many variables contribute collectively to the variation captured by PC-1. However, a sharp decline around the ~1000 cm^{-1} region suggests that this region has a distinct variation and may act in an opposite or less contributing manner compared to other variables. This region is typically associated with SiO_2, Si–O lattice stretching vibrations (silicates, quartz), indicating that these components play a critical role in differentiating the samples. Overall, the loading plot confirms that while most spectral regions contribute positively, specific bands, particularly around 1000 cm^{-1} are highly influential in

distinguishing between samples, making them key markers for forensic interpretation and comparison.

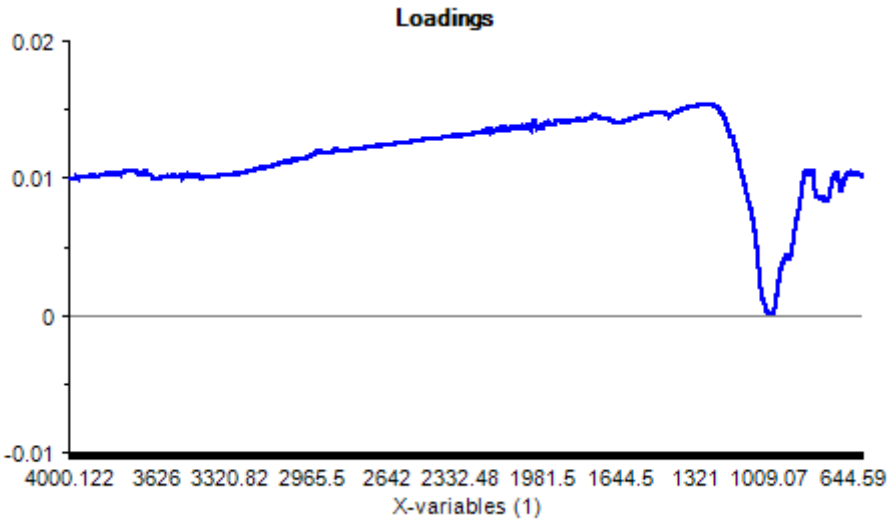


Fig. 5 Loading plot of PCA

Fig. 6 shows the zoom in view of the desegregated cluster consisting of samples 7-15. These clustered samples belong to the same geographical area of Keralam (7-Angamaly; 08-CUSAT; 09-Ernakulam; 10- Kannur; 11- Pala, Kottayam; 12- Palakkad; 13- Pali Malli; 14-Perumbavoor; 15- Thrissur). PCA was further applied in second step using these samples to study the discrimination among these samples.

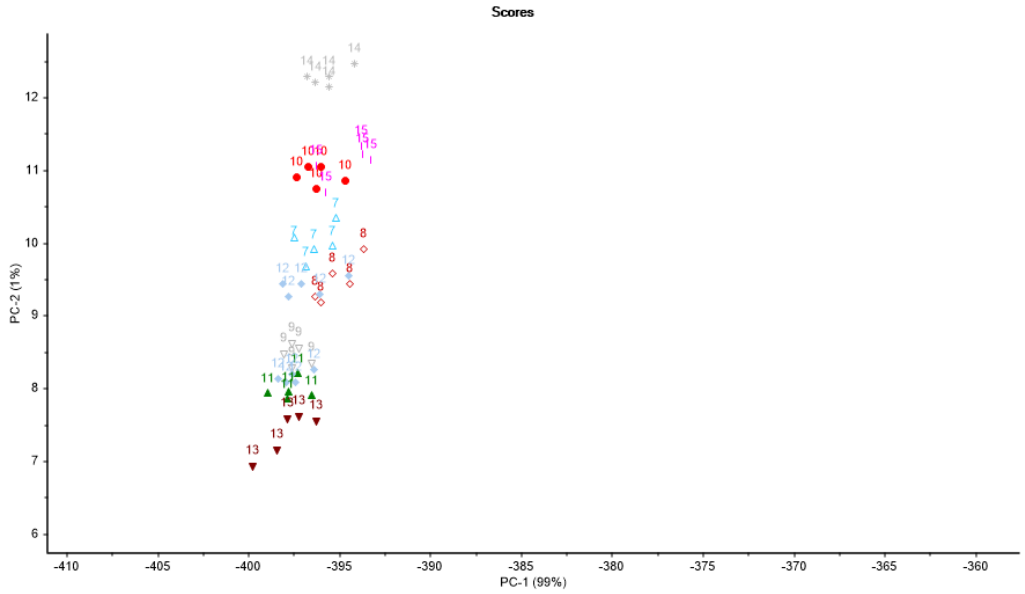


Fig. 6 Zoom-in View of PCA plot to separate the cluster of samples 7-15

The PCA score plot (Fig.7.) showed the distinct clustering and separation patterns among the soil samples collected from different geographical regions. Sample number 7 (Angamaly), Sample number 10 (Kannur), sample number 11 (Pala, Kottayam), and sample number 14 (Perumbavoor) exhibited clear separation from the remaining samples, indicating distinct chemical compositions. The divergence observed between these samples, despite their proximity in geographic origin, strongly implies that they exhibit distinct soil properties, which are probably a consequence of variations in local geological structures, prevailing environmental circumstances, or human-induced activities.

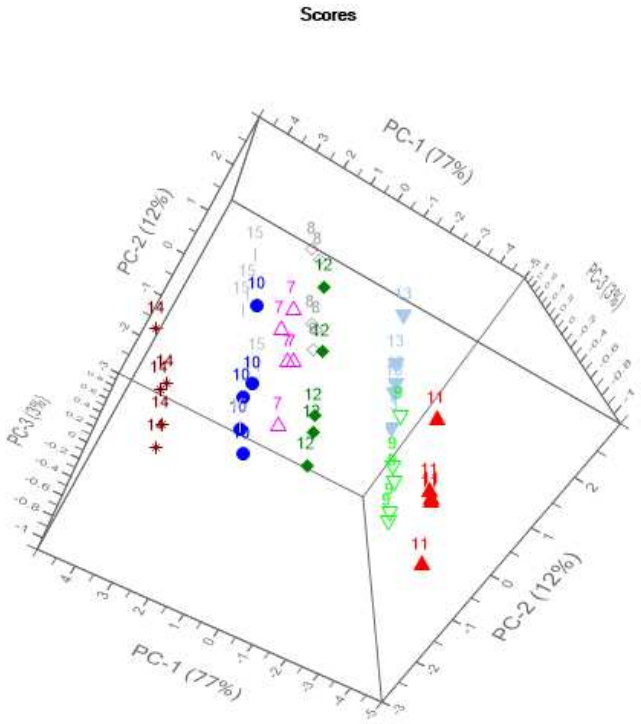


Fig.7. PCA of Sample no 7-15

Sample 15 (Thrissur) was placed closely with Sample 10 (Kannur), indicating misclassification of few samples. This suggests that both locations may share similar spectral profiles and compositional features, possibly due to comparable soil types or regional proximity within Keralam. Sample 8 (CUSAT, Cochin University Area) showed similar misclassification with Sample 12 (Palakkad), implying similarity in their chemical signatures, despite being from different micro-regions. Samples 9 (Ernakulam) and 13 (Pali Malli) were found to be closer in the PCA space. The PCA model achieved an overall classification accuracy of 86.11%, indicating good discriminatory capability among the analyzed classes. However, some misclassification was observed, particularly for Class 8, Class 9, and Class 15 samples (Table 4)

Table 4: Classification and Misclassification of classes using PCA model

Actual Class	Total Samples	Correctly Classified	Accuracy (%)	Misclassified	Misclassified As	Misclassification (%)
Class 1	6	6	100%	0	–	0%
Class 2	6	6	100%	0	–	0%
Class 3	6	6	100%	0	–	0%
Class 4	6	6	100%	0	–	0%
Class 5	6	6	100%	0	–	0%
Class 6	6	6	100%	0	–	0%
Class 8	6	2	33.33%	4	Class 12	66.67%
Class 9	6	4	66.67%	2	Class 13	33.33%
Class 10	6	6	100%	0	–	0%
Class 11	6	6	100%	0	–	0%
Class 14	6	6	100%	0	–	0%
Class 15	6	2	33.33%	4	Class 10	66.67%

3.4 Soil discrimination using PLS-DA model

The PLS-DA score plot illustrates the supervised classification of the samples based on their spectral characteristics. The explained variance plot of the PLS-DA (Fig. 8) showed that there are 3 factors which are significant for the discrimination between the soil samples.

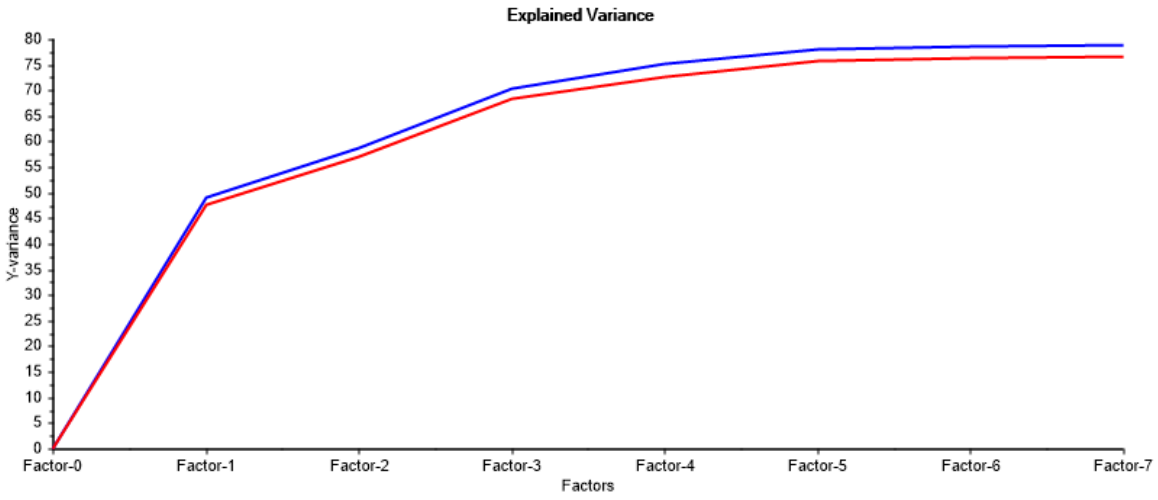


Fig.8. Explained variance plot

The PLS-DA model (Fig. 9) shows that Factor-1 explains 99% of the variance, indicating that it is the primary component responsible for distinguishing between the different sample classes, while Factor-2 (1%) and Factor-3 (~0%) contribute minimally to the overall

variation. The samples are grouped into well-defined and distinct clusters, except for a few sample groups.

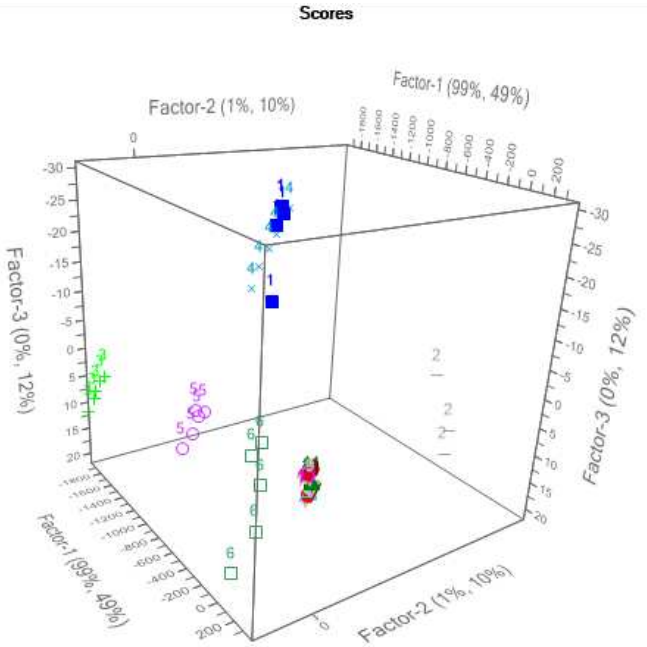


Fig.9 3D-PLSDA plot for the discrimination of soil samples

The results indicate that samples number 1, 2, 3, 4, 5, and 6 were clearly separated into distinct clusters, demonstrating that the PLS-DA model effectively captured the variability among these classes, like PCA. However, sample 1 (Bangoru Falls) showed proximity to the cluster of samples 4 (Mohan Garden, New Delhi), indicating partial overlap or misclassification. This suggests that these samples may share similar chemical or spectral features, leading to reduced discrimination between them.

In addition, the remaining samples (7–15) formed a tight cluster, like the observation in PCA. This indicates that these samples possess high compositional similarity, likely due to their common geographical origin (Keralam region), which limits their separation even under supervised analysis.

To address this limitation, a separate PLS-DA model was further developed for the tightly clustered group (samples 7–15). This targeted modelling approach enhances the resolution of subtle differences within this group, allowing improved classification and discrimination. Calibration and validation values of the PLS-DA model are shown in Table 5. Predicated vs reference plot is given in Fig. 10

Table 5 Calibration and validation using PLS-DA model

Parameter	**Blue (Calibration)**	**Red (Validation)**
Slope	0.78	0.76
Offset	1.84	2.11
RMSE	1.98	2.10
R^2	0.78	0.76

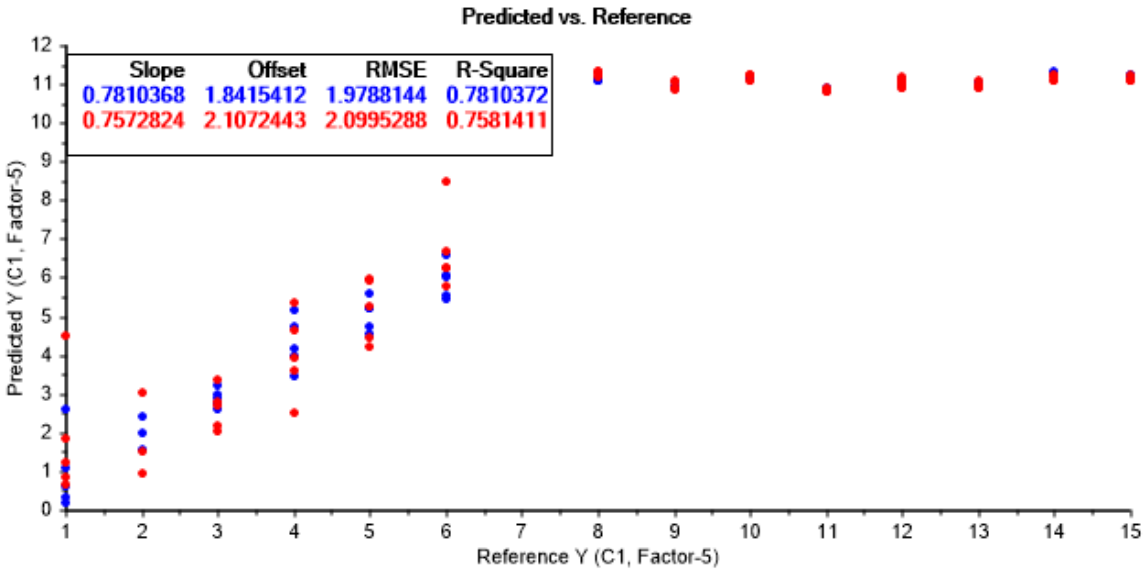


Fig.10 Predicted vs Reference plot to discriminate the Soil samples

Fig. 10 shows the Predicted Vs Reference PLS-DA plot to discriminate soil samples collected from different geographical regions. The PLS-DA Predicted vs. Reference plot

indicates that the R^2 values (0.78) suggest that the model explains a reasonable but not strong proportion of variance. slope values (<1) indicate underfitting, meaning the model compresses predictions and does not fully capture class separation. Additionally, the positive offset (1.8) reflects a systematic bias, where predicted values are consistently higher than actual values. The relatively high RMSE (1.9) further confirms the presence of prediction errors and possible misclassification among samples. Overall, the model shows acceptable but not robust discrimination, suggesting that some classes are not well separated and may require further optimisation or additional variables for improved accuracy. Further to improve the accuracy, the sample group 7-15 were used to discriminate these closely assembled samples.

However, the model was further developed with a single cluster of various samples (Fig. 11) . Results showed that all samples were separated into separate clusters; however, only sample number 7 (Angamaly) showed closeness with sample number 8 (CUSAT), and a few samples in group 12 (Palakkad) were misclassified with group 10 (Kannur). The high R-squared value was obtained for the separation of samples collected from keralam region and thus indicates the robustness of the model.

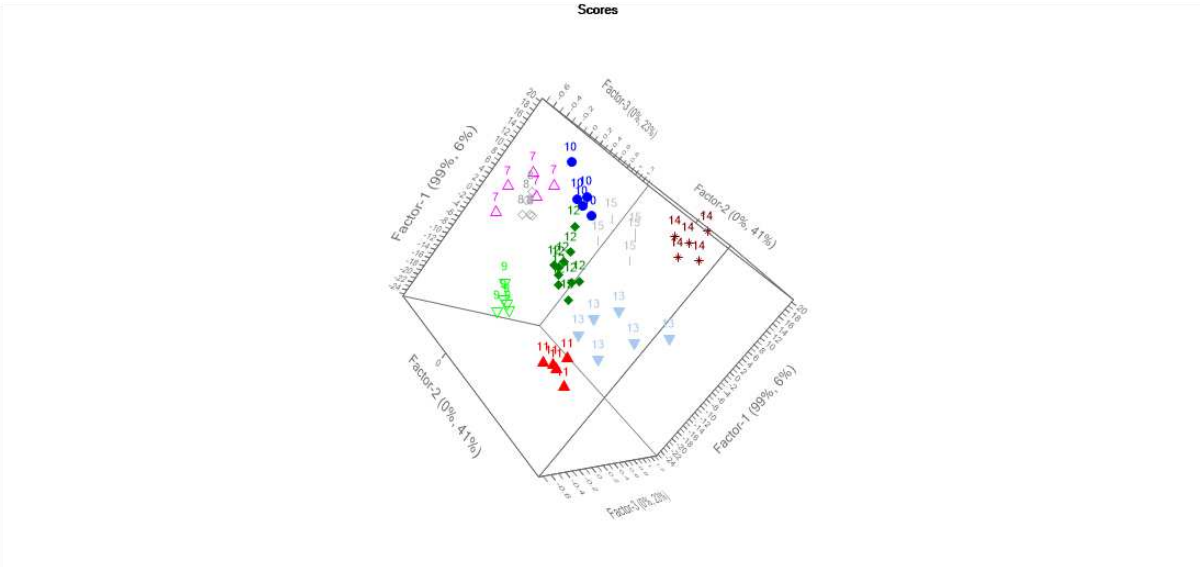


Fig. 11 PLS-DA plot of 7-15 samples

The PLS-DA predicted vs. reference plot demonstrated good discrimination for samples from the Keralam region. The calibration model (blue points) (fig. 12) shows a strong linear relationship with a slope close to 1 (0.9838) and a high R^2 value (0.9838), indicating excellent agreement between predicted and actual class values and reliable model fitting. The low RMSE (0.308) further confirms minimal prediction error during calibration. In contrast, the validation model (red points) shows a comparatively lower slope (0.7243) and R^2 (0.8342), along with a higher RMSE (1.0209), suggesting some reduction in predictive accuracy when applied to unknown samples. Despite this, the clustering pattern across class values (7–15) remains reasonably well separated, indicating that the model is robust in discriminating the Kerala region samples, though with slight overlap and variability in certain classes.

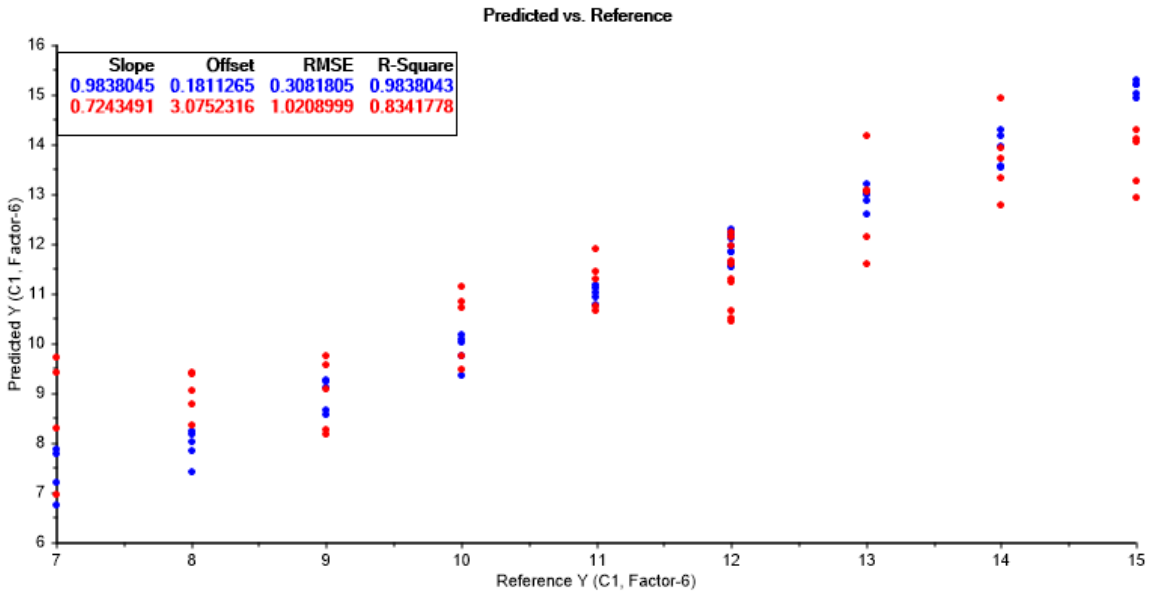


Fig. 12 - Predicted Vs Reference plot (Sample 7-15)

The developed model showed good discriminatory power for both principal component analysis (PCA) and partial least square – discriminant analysis (PLS-DA) methods. Also, the

developed models show low RMSE values and high R-square values, which prove the robustness of the developed models.

Table 6 showed that the correct classification (100% accuracy) was observed for Classes 2, 3, 4, 5, 6, 8, 9, 10, 11, 13, 14, and 15. However, all samples of Class 1 were misclassified as Class 4, and all samples of Class 7 were misclassified as Class 8, indicating spectral similarity between these respective classes. Additionally, one sample from Class 12 was misclassified as Class 10, while the remaining samples were correctly classified. Overall, the PLS-DA model achieved an accuracy of 85.56%, demonstrating good discriminatory capability among the majority of the sample classes.

Table:6 Classification and Misclassification of classes using PCA model

Actual Class	Total Samples	Correctly Classified	Misclassified	Misclassified As	Accuracy (%)	Misclassification (%)
Class 1	6	0	6	Class 4	0.00	100.00
Class 2	6	6	0	–	100.00	0.00
Class 3	6	6	0	–	100.00	0.00
Class 4	6	6	0	–	100.00	0.00
Class 5	6	6	0	–	100.00	0.00
Class 6	6	6	0	–	100.00	0.00
Class 7	6	0	6	Class 8	0.00	100.00
Class 8	6	6	0	–	100.00	0.00
Class 9	6	6	0	–	100.00	0.00
Class 10	6	6	0	–	100.00	0.00
Class 11	6	6	0	–	100.00	0.00
Class 12	6	5	1	Class 10	83.33	16.67
Class 13	6	6	0	–	100.00	0.00
Class 14	6	6	0	–	100.00	0.00
Class 15	6	6	0	–	100.00	0.00

Conclusions

Forensic soil analysis is an integral part of criminal investigation which could provide important investigative leads by linking crime scene, accused, and victim. The conventional methods used for the analysis of organic components of soil are often costly, chronophagic, and require extensive sample preparation. Attenuated total internal reflection - Fourier transform infrared (ATR-

FTIR) spectroscopy overcomes these limitations by providing a cheaper, faster, and reliable alternative which is exceptionally effective in discriminating samples. However, it is imperative to combine the spectroscopic results with the chemometric tools such as principal component analysis (PCA) and partial least square – discriminant analysis (PLS-DA) to uncover hidden patterns which can be used for accurately identifying the geographic provenance of the soil samples. The chemometric models demonstrated strong classification performance for the analyzed samples. The PCA model achieved a classification accuracy of 86.11%, indicating good separation among the classes, while the supervised PLS-DA model showed a classification accuracy of 85.56%, confirming its effectiveness in discriminating the sample groups despite minor misclassifications. The few misclassifications observed were mainly associated with samples from geographically closer regions of Kerala, possibly due to similarities in their chemical composition. The chemometric model developed in the present study shows high discrimination for both the chemometric tools along with high R-square value and low RMSE value, which establish the robustness of the developed model. Therefore, it can be effectively used in the real cases.

References

[1] Gross HGA. *Handbuch Für Untersuchungsrichter*. Schweitzer Verlag; 1908.

[2] Murray RC. Forensic examination of soils. *Forensic chemistry handbook*. Published online 2012:109-130.

[3] Wilding L. Factors of soil formation: contributions to pedology. *Factors of soil formation: a fiftieth anniversary retrospective*. 1994;33:15-30.

[4] Marumo Y. Forensic examination of soil evidence. *Japanese Journal of science and technology for identification*. 2003;7(2):95-111.

[5] FitzpatRick RW. Nature, distribution, and origin of soil materials in the forensic comparison of soils. In: *Soil Analysis in Forensic Taphonomy*. CRC Press; 2008:13-40.

[6] Fitzpatrick R. Forensic earth science: Getting the dirt on crime. e-Science Faculty of Sciences. *The University of Adelaide*. 2012;2:5-11.

[7] Fitzpatrick RW. Soil: forensic analysis. *Wiley encyclopedia of forensic science Chichester: John Wiley & Sons*. Published online 2009:2377-2388.

[8] Locard E. Locard's Exchange Principle. Published online 2008.

[9] Sangwan P, Nain T, Singal K, Hooda N, Sharma N. Soil as a tool of revelation in forensic science: a review. *Analytical methods*. 2020;12(43):5150-5159.

[10] Otto A, Simpson MJ. Analysis of soil organic matter biomarkers by sequential chemical degradation and gas chromatography–mass spectrometry. *Journal of Separation Science*. 2007;30(2):272-282.

[11] Santos F, Galceran M. The application of gas chromatography to environmental analysis. *TrAC Trends in Analytical Chemistry*. 2002;21(9-10):672-685.

[12] Chesters G, Konrad J, Schrag B, Everett L. Gas chromatography: Techniques and uses in soil, plant, and water analysis. *Instrumental Methods for Analysis of Soils and Plant Tissue*. Published online 1971:129-183.

[13] Goyal SS. Use of high performance liquid chromatography for soil and plant analysis. *Communications in soil science and plant analysis*. 2002;33(15-18):2617-2641.

[14] McCulloch G, Morgan R, Bull P. High performance liquid chromatography as a valuable tool for geoforensic soil analysis. *Australian Journal of Forensic Sciences*. 2017;49(4):421-448.

[15] Baziramakenga R, Simard R, Leroux G. Determination of organic acids in soil extracts by ion chromatography. *Soil Biology and Biochemistry*. 1995;27(3):349-356.

[16] Michalski R, Muntean E, Pecyna-Utylska P, Kernert J. Ion chromatography-an advantageous technique in soil analysis. *ProEnvironment*. 2019;12:82-88.

[17] Xing Z, Du C, Zeng Y, Ma F, Zhou J. Characterizing typical farmland soils in China using Raman spectroscopy. *Geoderma*. 2016;268:147-155.

[18] Kammrath BW, Koutrakos A, Castillo J, Langley C, Huck-Jones D. Morphologically-directed Raman spectroscopy for forensic soil analysis. *Forensic Science International*. 2018;285:e25-e33.

[19] Maiwald M, Sowoidnich K, Sumpf B. Portable shifted excitation Raman difference spectroscopy for on-site soil analysis. *J Raman Spectroscopy*. 2022;53(9):1560-1570. doi:10.1002/jrs.6400

[20] Malley DF, Yesmin L, Wray D, Edwards S. Application of near-infrared spectroscopy in analysis of soil mineral nutrients. *Communications in Soil Science and Plant Analysis*. 1999;30(7-8):999-1012. doi:10.1080/00103629909370263

[21] Ellerbrock RH, Höhn A, Gerke HH. Characterization of soil organic matter from a sandy soil in relation to management practice using FT-IR spectroscopy. *Plant and Soil*. 1999;213(1):55-61. doi:10.1023/A:1004511714538

[22] Tüzen M. Determination of heavy metals in soil, mushroom and plant samples by atomic absorption spectrometry. *Microchemical Journal*. 2003;74(3):289-297. doi:10.1016/S0026-265X(03)00035-3

[23] Croffie MET, Williams PN, Fenton O, Fenelon A, Metzger K, Daly K. Optimising Sample Preparation and Calibrations in EDXRF for Quantitative Soil Analysis. *Agronomy*. 2020;10(9):1309. doi:10.3390/agronomy10091309

[24] Singh V, Agrawal HM. Qualitative soil mineral analysis by EDXRF, XRD and AAS probes. *Radiation Physics and Chemistry*. 2012;81(12):1796-1803. doi:10.1016/j.radphyschem.2012.07.002

[25] Fernández ZH, dos Santos Júnior JA, dos Santos Amaral R, Alvarez JRE, da Silva EB, De França EJ, et al. EDXRF as an alternative method for multielement analysis of tropical soils and sediments. *Environmental Monitoring and Assessment*. 2017;189(9):447. doi:10.1007/s10661-017-6162-5

[26] Moor C, Lymberopoulou T, Dietrich VJ. Determination of Heavy Metals in Soils, Sediments and Geological Materials by ICP-AES and ICP-MS. *Microchimica Acta*. 2001;136(3):123-128. doi:10.1007/s006040170041

[27] Payá-pérez A, Sala J, Mousty F. Comparison of ICP-AES and ICP-MS for the Analysis of Trace Elements in Soil Extracts. *International Journal of Environmental Analytical Chemistry*. 1993;51(1-4):223-230. doi:10.1080/03067319308027628

[28] Yamasaki S, Tamura Y. Total analysis of soils by ICP-AES and ICP-MS (Mass spectrometry). *Communications in Soil Science and Plant Analysis*. 1990;21(17-18):2017-2028. doi:10.1080/00103629009368356

[29] Assefa T, Mathur S. The Role of Soil as Trace Evidence in Forensic Science: Methodologies and Applications. *Indian Journal of Forensic Medicine & Toxicology*. 2024;19(1):38-44. doi:10.37506/arsd9w59

[30] Skoog DA, West DM, Holler FJ, Crouch SR. *Fundamentals of Analytical Chemistry*. Brooks/Cole Cengage Learning; 2014.

[31] Cox RJ, Peterson HL, Young J, Cusik C, Espinoza EO. The forensic analysis of soil organic by FTIR. *Forensic Science International*. 2000;108(2):107-116. doi:10.1016/S0379-0738(99)00203-0

[32] Force J, Elkins KM. FTIR spectroscopic analysis of soil in forensic science. *Frontiers in Analytical Science*. 2025;Volume 5-2025. https://www.frontiersin.org/journals/analytical-science/articles/10.3389/frans.2025.1716867

[33] Xu X, Du C, Ma F, Shen Y, Zhou J. Forensic soil analysis using laser-induced breakdown spectroscopy (LIBS) and Fourier transform infrared total attenuated

reflectance spectroscopy (FTIR-ATR): Principles and case studies. *Forensic Science International*. 2020;310:110222. doi:10.1016/j.forsciint.2020.110222

[34] Thabit FN, Negim OIA, AbdelRahman MAE, Scopa A, Moursy ARA. Using Various Models for Predicting Soil Organic Carbon Based on DRIFT-FTIR and Chemical Analysis. *Soil Systems*. 2024;8(1):22. doi:10.3390/soilsystems8010022

[35] Mahmoudi MR, Heydari MH, Qasem SN, Mosavi A, Band SS. Principal component analysis to study the relations between the spread rates of COVID-19 in high risks countries. *Alexandria Engineering Journal*. 2021;60(1):457-464. doi:https://doi.org/10.1016/j.aej.2020.09.013

[36] Kumar R, Sharma V. Chemometrics in forensic science. *TrAC Trends in Analytical Chemistry*. 2018;105:191-201. doi:https://doi.org/10.1016/j.trac.2018.05.010