

Rapid Identification and Quantification of Adulteration in Methyl Eugenol using Raman Spectroscopy Coupled with Multivariate Data Analysis

Muntaha Anwar,[#] Gull Rimsha,[#] Muhammad Irfan Majeed,^{*} Najah Alwadie,^{*} Haq Nawaz,^{*} Muhammad Zeeshan Majeed, Nosheen Rashid, Fareeha Zafar, Ali Kamran, Muhammad Wasim, Nasir Mehmood, Ifra Shabbir, and Muhammad Imran



Cite This: *ACS Omega* 2024, 9, 7545–7553



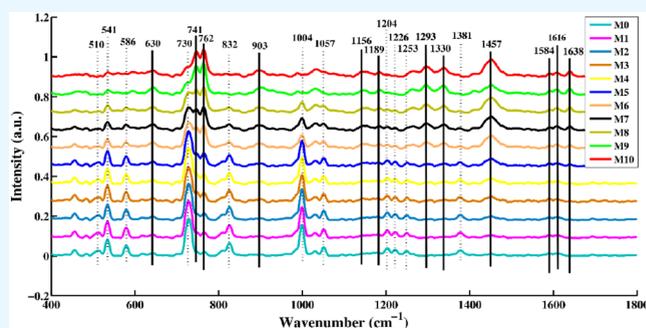
Read Online

ACCESS |

Metrics & More

Article Recommendations

ABSTRACT: Identification of adulterants in commercial samples of methyl eugenol is necessary because it is a botanical insecticide, a tephritid male attractant lure that is used to attract and kill invasive pests such as oriental fruit flies and melon flies on crops. In this study, Raman spectroscopy was used to qualitatively and quantitatively assess commercial methyl eugenol along with adulterants. For this purpose, commercial methyl eugenol was adulterated with different concentrations of xylene. The Raman spectral features of methyl eugenol and xylene in liquid formulations were examined, and Raman peaks were identified as associated with the methyl eugenol and adulterant. Principal component analysis (PCA) and partial least-squares regression analysis (PLSR) have been used to qualitatively and quantitatively analyze the Raman spectral features. PCA was applied to differentiate Raman spectral data for various concentrations of methyl eugenol and xylene. Additionally, PLSR has been used to develop a predictive model to observe a quantitative relationship between various concentrations of adulterated methyl eugenol and their Raman spectral data sets. The root-mean-square errors of calibration and prediction were calculated using this model, and the results were found to be 1.90 and 3.86, respectively. The goodness of fit of the PLSR model is found to be 0.99. The proposed approach showed excellent potential for the rapid, quantitative detection of adulterants in methyl eugenol, and it may be applied to the analysis of a range of pesticide products.



1. INTRODUCTION

Methyl eugenol (ME) is a botanical insecticide,¹ and it belongs to a class of phenylpropanoids² with many synonyms including eugenol methyl ether, allylveratrol, 3,4-dimethoxy-allylbenzene, and *O*-methyl eugenol.³ It is a colorless to pale yellow liquid with a molecular formula of C₁₁H₁₄O₂.⁴ It is a highly potent kairomone lure⁵ that occurs naturally in a wide range of aromatic plants, particularly as a constituent of natural essential oils.⁶ It is a powerful inhibitor of the acetylcholinesterase enzyme, which is responsible for hydrolyzing the neurotransmitter acetylcholine that ultimately causes paralysis in insects.⁷ In the USA in 2006, methyl eugenol was authorized to be used as an active pesticide ingredient⁸ and is usually considered safe by the United States Food and Drug Administration.⁵ It is used as an insect attractant in pesticide formulations⁹ and has antibacterial, antifungal, and antinematodal properties.¹⁰ Hence, it has a high economic value.¹¹

Pesticides are available in several “formulations” and are marketed under different trademarks.¹² The active ingredient

and inert substances are combined for the preparation of these formulations.¹³ Each active ingredient is often available in a variety of commercial formulations, and it is well-known that inert ingredients can significantly alter the availability and toxicity of the active ingredient.¹⁴ Inert ingredients are used in pesticide formulations for a wide range of purposes, including preservatives, surfactants, and solvents.¹³ Xylene is the most common solvent used as an inert ingredient in many pesticide products.¹⁵ Xylene may be used as an adulterant in pesticides. Therefore, it is necessary to determine the quality and quantity of the adulterants in pesticide formulations. In the present work, xylene has been used as an adulterant in commercial

Received: August 25, 2023

Revised: January 9, 2024

Accepted: January 26, 2024

Published: February 8, 2024



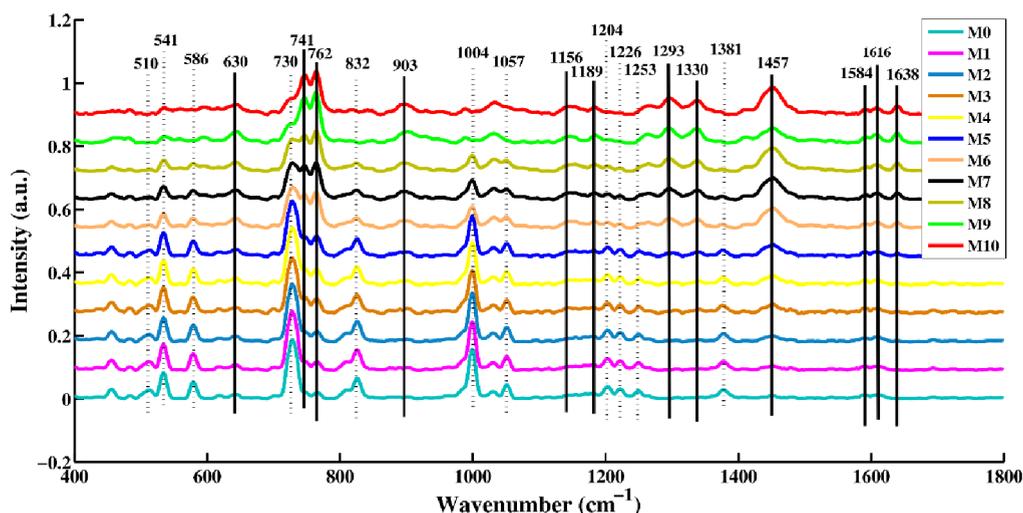


Figure 1. Mean Raman spectra of all adulterated methyl eugenol samples (M0; pure xylene; M1–M9; different concentrations of xylene and methyl eugenol; M10; commercial methyl eugenol).

methyl eugenol to make different formulations to check the quality and quantity of the adulterant and methyl eugenol.

Adulterants are substances that are consciously added to more expensive products to increase their actual quantities, lower their manufacturing costs, or for other fraudulent objectives.¹⁶ Pesticide adulteration may occur during the stages of production, storage, delivery, or circulation.¹⁷ The adulteration of pesticides is extremely detrimental to agricultural production and may also be hazardous to the environment.¹⁸ The detection and quantification of adulterants in pesticide products is regarded as a usual part of quality control operations.¹⁹ To achieve this goal and ensure the quality of plant protection products (i.e., pesticides), different chromatographic and spectroscopic techniques have been employed.²⁰ Many studies have shown that different conventional analytical techniques have been used for analyzing the concentration of adulterants in pesticide formulations, such as liquid chromatography mass spectrometry (LC-MS) and gas chromatography mass spectrometry (GC-MS) for the determination of adulterants in pesticide formulations,²¹ high-performance liquid chromatography and gas chromatography for the detection of active ingredients and adulterants present in plant protection products,²² ultra-performance liquid chromatography–mass spectrometry (UPLC-MS) and nuclear magnetic resonance (NMR) spectroscopy for the analysis of adulterants contained in pesticide formulations,²³ liquid chromatography–tandem mass spectrometry for the quantification of contaminants and adulterants present in pesticide formulations,²⁴ high-performance liquid chromatography along with diode-array detector (HPLC-DAD) for the simultaneous identification of adulterants and active ingredients in pesticide formulations,²⁵ and Fourier transform infrared (FTIR) spectroscopy to detect the adulterants present in pesticide products.²⁶

These techniques are costly, time-consuming, and require sample preparation.²⁷ Hence, there is a need to develop a rapid, precise, nondestructive, efficient, and reliable technique to detect the adulterants present in pesticide formulations. Raman spectroscopy is a quick method and efficient analytical tool that offers great advantages such as being nondestructive, less expensive, and providing ultrasensitive identification down to a single molecular level.²⁸ This technique has been

employed for various applications including drug analysis,^{29,30} disease diagnosis,^{31,32} and characterization of organometallic compounds.³³ Raman spectroscopy has rarely been used to directly identify the active ingredients in various commercial pesticide formulations.³⁴ There is no literature that was published to check the quality and quantity of commercial methyl eugenol along with adulterant (xylene). In this research, liquid samples are prepared in various concentrations by mixing commercial methyl eugenol and xylene in order to access the quality and quantity of methyl eugenol along with adulterants in pesticide samples by applying Raman spectroscopy, as well as multivariate data analysis methods such as principal component analysis (PCA) and partial least-squares regression analysis (PLSR).

Although xylene is an inert substance, as an adulterant it can affect the activity of pesticides to some extent by diluting the actual concentration of pesticide. When xylene is added as an adulterant, it actually lowers the claimed or expected efficiency of that pesticide. There is existing literature that discusses various techniques for analyzing the concentration of xylene in pesticide formulations. For instance, rapid detection of xylene, a common pesticide adulterant, can be successfully quantitated using pulse heating–gas chromatography–mass spectrometry (Py-GC-MS) in blood and urine.³⁵ Gas chromatography–mass spectrometry (GC-MS) has also been employed to determine the presence of six pesticide adulterants, including toluene, *p*-xylene, *o*-xylene, *m*-xylene, *N,N*-dimethylformamide (DMF), and dimethyl sulfoxide (DMSO).³⁶ Furthermore, solid-phase microextraction has been applied to analyze various organic compounds, including benzene, toluene, ethylbenzene, and xylene.³⁷

However, it is important to note that there is no existing literature to check the quality and quantity of commercial methyl eugenol along with the adulterant (xylene) by using Raman spectroscopy. Therefore, the primary objective of this study is to use Raman spectroscopy to quantify the presence of xylene as an adulterant as it is a rapid, precise, nondestructive, efficient, and reliable technique when compared to other conventional methods.

2. RESULTS AND DISCUSSION

2.1. MEAN RAMAN SPECTRA

The mean Raman spectra of all of the samples with various concentrations of methyl eugenol and xylene (an adulterant) are displayed in Figure 1. The main differences in the Raman spectra are depicted by vertical lines, where the spectral features of xylene and commercial methyl eugenol are shown by dotted and solid lines, respectively. A list of Raman peak assignments obtained from the literature with respective references is shown in Table 1.

Table 1. Raman Peaks of all Samples of Pesticide (Methyl Eugenol) Assigned from the Literature with References

| Raman shift (cm ⁻¹) | peak assignments | references |
|---------------------------------|---|------------|
| 510 | <i>o</i> -xylene: ring C–C–C out of plane bending | 38 |
| 541 | <i>m</i> -xylene: ring breathing | 38 |
| 586 | <i>o</i> -xylene: ring breathing | 38 |
| 630 | ring deformation | 28 |
| 730 | CCC out of plane bending in <i>m</i> -xylene ring | 38 |
| 741 | ring deformation | 39 |
| 762 | ring deformation | 39 |
| 832 | out-of-plane bending of CH + ring bending of <i>p</i> -xylene | 38 |
| 903 | back-bone stretching in C–C bond | 28 |
| 1004 | <i>m</i> -xylene: ring bending | 38 |
| 1057 | <i>o</i> -xylene: ring breathing | 38 |
| 1156 | C–C bond stretching | 28 |
| 1189 | ring breathing | 39 |
| 1204 | In plane bending of CH bond + ring bending of <i>m</i> -xylene | 38 |
| 1226 | Stretching vibrations of CH ₃ + C–H bond in plane bending + ring bending of <i>o</i> -xylene | 38 |
| 1253 | C=C stretching + ring bending of <i>m</i> -xylene | 38 |
| 1293 | twisting vibration of CH ₂ | 39 |
| 1330 | CH ₂ twisting vibration | 39 |
| 1381 | C–H bond stretching in meta-xylene ring | 38 |
| 1457 | deformation of CH ₂ and CH ₃ | 40 |
| 1584 | C=C (aromatic) bond stretching | 39 |
| 1616 | stretching vibration in conjugated C=C bond | 39 |
| 1638 | C=CH bond wagging vibration | 41 |

The changes in spectral characteristics are associated with an increase in the concentration of methyl eugenol and a decrease in the xylene concentration (an adulterant). The significant Raman peaks of commercial methyl eugenol were observed at 630, 741, 762, 903, 1156, 1089, 1293, 1330, 1457, 1584, 1616, and 1638 cm⁻¹. In Figure 1, it is shown that the intensities of the methyl eugenol peaks gradually increased as the concentration of methyl eugenol increased. The deformation of the ring is related to the moderate Raman peak that appeared at 630 cm⁻¹. The strong peaks of methyl eugenol observed at 741 and 762 cm⁻¹ are related to ring deformation. The backbone stretching of the C–C bond is associated with a low Raman peak at 903 cm⁻¹, and other weak intensity peaks at 1156 and 1089 cm⁻¹ correspond to the C–C stretching and ring breathing, respectively. The moderate intensity peak observed at 1293 cm⁻¹ represents the CH₂ twisting vibration. In addition, the twisting vibration of CH₂ and the deformation of CH₂ and CH₃ are connected with moderate intensity peaks at 1330 and 1457 cm⁻¹. The low-intensity peaks that appeared at 1584 and 1616 cm⁻¹ show the C=C (aromatic) stretching and conjugated C=C bond stretching. The wagging vibration of C=CH is associated with a moderate intensity peak at 1638 cm⁻¹.

The Raman peaks of xylene are observed at 510, 541, 586, 730, 832, 1004, 1057, 1204, 1226, 1253, and 1381 cm⁻¹. The out-of-plane bending of the C–C–C in the ortho xylene ring is related to the weak Raman peak at 510 cm⁻¹. The ring breathing of meta xylene is indicated by a moderate Raman peak at 541 cm⁻¹ and another moderate intensity peak at 586 cm⁻¹ representing the ring breathing of ortho xylene. The out of plane bending of the C–C–C bond in the meta xylene ring is represented by the high-intensity Raman peak seen at 730 cm⁻¹. The moderate Raman peak found at 832 cm⁻¹ is connected with CH out-of-plane bending and ring bending of para xylene.

A significant peak observed at 1004 cm⁻¹ is linked to the ring bending of *m*-xylene. The moderate intensity peak at 1057 cm⁻¹ indicates *o*-xylene ring breathing. Furthermore, a low-intensity band at 1204 cm⁻¹ is connected to the in-plane bending of C–H and meta-xylene ring bending. The weak Raman peak observed at 1226 cm⁻¹ represents the stretching vibrations of the CH₃ bond, in-plane bending of the CH bond,

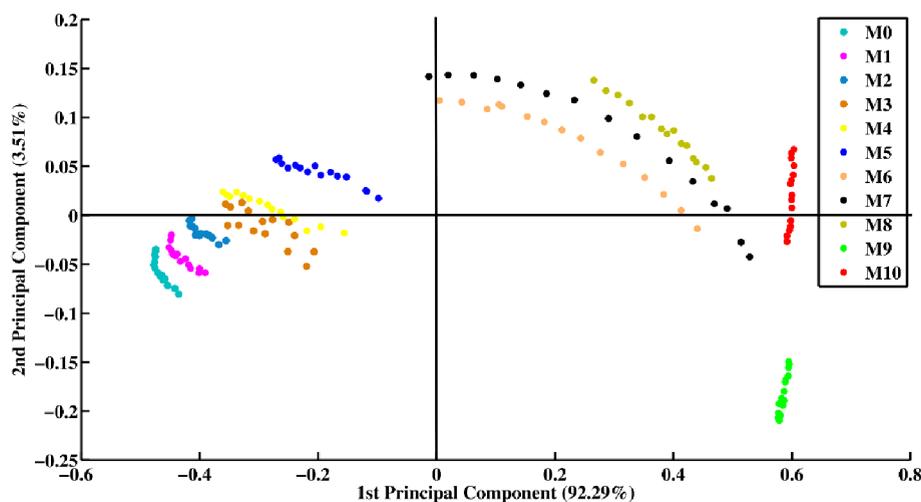


Figure 2. PCA scatter plot of pure methyl eugenol and samples of methyl eugenol adulterated with different concentrations of xylene.

and ring bending of ortho xylene. A weak intensity peak appeared at 1253 cm^{-1} linked with the stretching vibrations of C=C and the ring bending of meta xylene. The stretching vibration of the C–H bond in the meta xylene ring is associated with a moderate peak at 1381 cm^{-1} . However, it is possible to identify the xylene and methyl eugenol in a prepared formulation using the characteristic Raman peaks of xylene and methyl eugenol, which are useful for both qualitative and quantitative analysis.

2.2. PRINCIPAL COMPONENT ANALYSIS (PCA)

Figure 2 depicts the PCA scatter plot of all spectra acquired from various samples (M0–M10). It is demonstrated with clarity that the Raman spectral data sets of all samples with varying concentrations (M0–M10) are clearly differentiated and clustered individually, each represented by a separate color. However, each cluster indicates the spectral characteristics of a single concentration/formulation. Moreover, it separates all of the clusters of various amounts of adulterated methyl eugenol samples in order from M0 to M10, indicating 92.29% variability for PC1 and 3.51% variability in the data for PC2.

The pairwise PCA scatter plots of methyl eugenol and xylene are shown in Figure 3a. The PCA loadings for all data (M0–

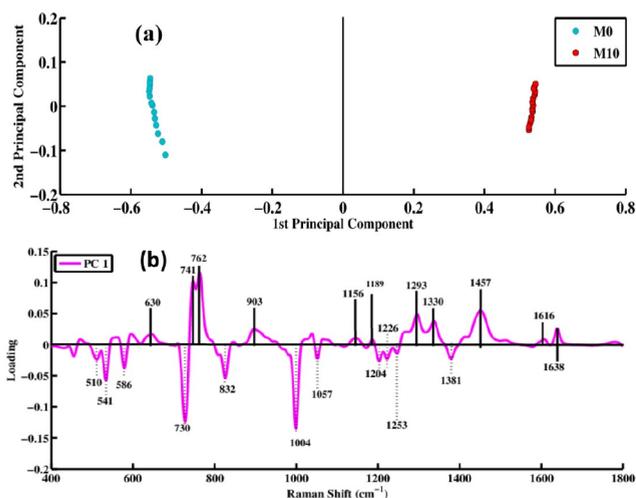


Figure 3. (a) pairwise PCA scatter plot of Raman spectra of pure xylene (M0) versus commercial methyl eugenol (M10). (b) PCA loadings of first principal component (PC-1) of Raman spectral data of xylene and commercial methyl eugenol.

M10) are displayed in Figure 3b. The Raman spectral characteristics of commercial methyl eugenol and xylene are represented by solid and dotted lines. The spectral properties of the mean plot Raman spectra are confirmed by these loadings. The spectral features of methyl eugenol appeared as positive loadings in PC-1 at 630 cm^{-1} (ring deformation), 741 cm^{-1} (ring deformation), 762 cm^{-1} (deformation of the ring), 903 cm^{-1} (back-bone stretching in C–C bond), 1156 cm^{-1} (C–C bond stretching), 1189 cm^{-1} (ring breathing), 1293 cm^{-1} (CH_2 twisting vibration), 1330 cm^{-1} (twisting vibration of CH_2), 1457 cm^{-1} (Deformation of CH_2 and CH_3), 1584 cm^{-1} (aromatic C=C bond stretching), 1616 cm^{-1} (conjugated C=C stretching), and 1638 cm^{-1} (wagging vibration of C=CH bond).

However, the negative loadings in PC-1 are associated with the spectral features of xylene (an adulterant), and these peaks are observed at 510 cm^{-1} (ortho-xylene: out of plane bending of C–C–C bond), 541 cm^{-1} (ring breathing of meta xylene), 586 cm^{-1} (ortho-xylene: ring breathing), 730 cm^{-1} (out-of-plane bending of C–C–C bond in meta xylene ring), 832 cm^{-1} (para-xylene: ring bending and out-of-plane bending of C–H bond), 1004 cm^{-1} (ring bending of m-xylene), 1057 cm^{-1} (ring breathing of ortho xylene), 1204 cm^{-1} (meta-xylene: ring bending and in-plane bending of CH bond), 1226 cm^{-1} (CH_3 stretching vibration + CH in-plane bending and ring bending of ortho xylene), 1253 cm^{-1} (C=C stretching vibration and the ring bending of meta xylene), and 1381 cm^{-1} (CH stretching in meta xylene ring). These positive and negative loadings confirm the observations of the mean Raman spectra presented in Figure 1. This shows that PCA is efficient in detecting and separating spectral data sets of various samples (M0–M10) with varying concentrations based on the distinctive Raman characteristic peaks of xylene and methyl eugenol.

The loadings of PC1 and PC2 are obtained through principal component analysis (PCA), a multivariate technique that identifies patterns in the data by capturing the directions of maximum variance. These loadings indicate how each data point, such as Raman shift, contributes to the variance observed in the data set.

In the case of Raman spectroscopy, PC loadings can be thought of as coefficients that relate specific Raman shifts or wavelengths to the overall variability in the data set. Positive loadings suggest that increases in Raman intensity at certain shifts contribute to the observed variance, while negative loadings suggest that decreases in intensity at those shifts are significant.

Principal component analysis provides significant structural information in Raman spectral data without a significant loss of information. In the scattered plot, PC1 is explained on the x -axis, while PC2 is explained on the y -axis for discrimination analysis of all samples of methyl eugenol prepared by mixing xylene at different concentrations. In the pairwise PCA scatter plot, we considered PC1 for comparison where M0 (xylene) appears on the negative side of x -axis and M10 (methyl eugenol) clustered on the positive side in first principal component analysis. So, we considered the loadings of PC1.

2.3. PARTIAL LEAST-SQUARES REGRESSION (PLSR) ANALYSIS

The PLSR model is used for the quantification/prediction of methyl eugenol along with xylene (an adulterant) in different formulations based on the response of Raman spectroscopy. To create a predictive model for the prediction of adulterated methyl eugenol in various samples, the PLSR model is developed using a standard sample containing a known concentration of xylene (Table 3). In the PLSR analysis, the SIMPLE algorithm was used by applying all 1499 wavenumbers as predictors (x -variables), and methyl eugenol and xylene (an adulterant) concentrations were employed as a response (y -variables). To avoid bias in the data analysis, all spectra of adulterated methyl eugenol with various concentrations were assembled into a matrix and randomly selected for modeling. For this analysis, two sets for total data, which include calibration and test sets, are employed, and samples were divided into 2 sets using a randomization of the data obtained from samples M0 to M10.

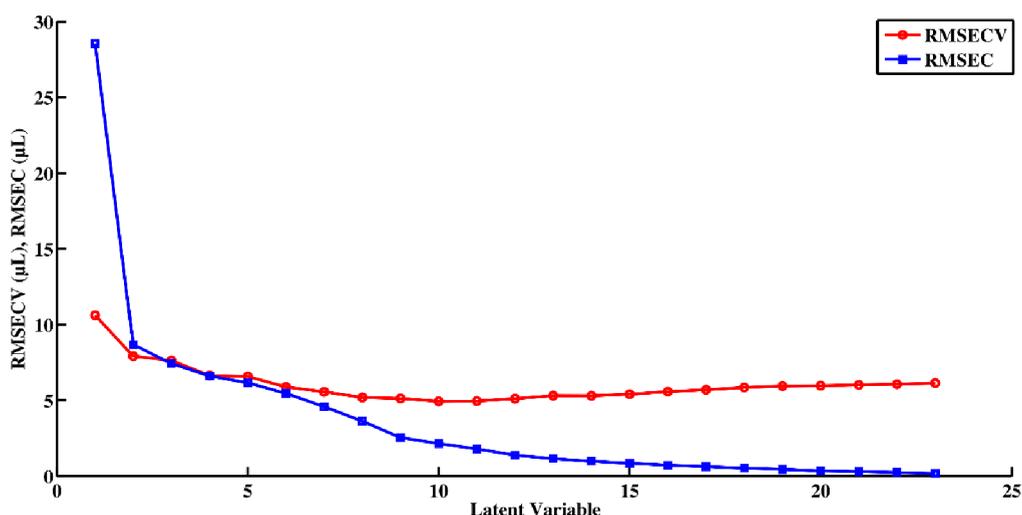


Figure 4. PLSR model for Raman spectral data of various adulterated samples of methyl eugenol along with xylene for the prediction of the optimal number of latent variables.

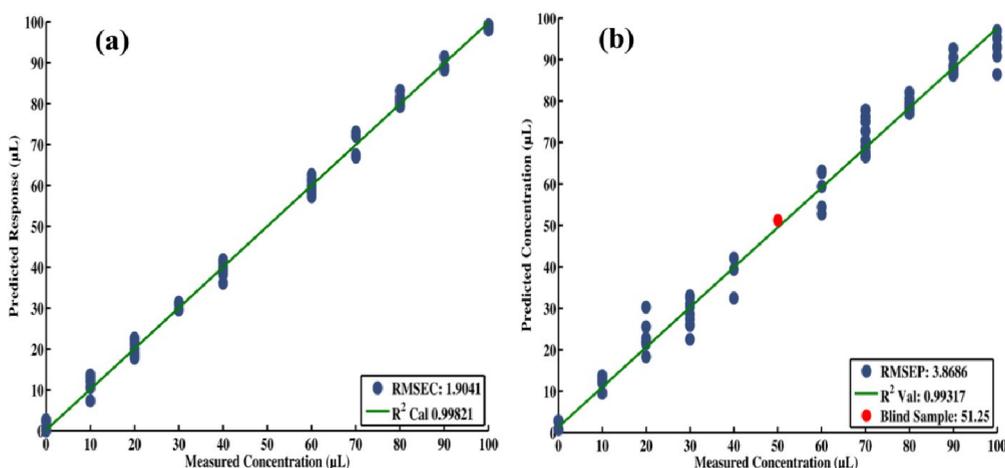


Figure 5. Performance of the PLSR model (calibration vs prediction) for Raman spectral data for the methyl eugenol.

Figure 4 shows the RMSECV of the first 23 latent variables. It is shown that in the above 6 latent variables the value tends to reduce at first before gradually increasing. The optimal number of latent variables is determined based on the lowest RMSECV value, whereas the risk of model overfitting should also be avoided. Thus, only 6 latent variables are selected, resulting in an RMSECV of 1.90. It is clear that above 6 latent variables, there is no remarkable improvement to the model.

2.4. ROOT-MEAN-SQUARE ERROR OF CROSS-VALIDATION

The PLSR model was created by choosing the optimal number of latent variables to evaluate the capability of this technique to predict the different concentrations of commercial methyl eugenol along with adulterant (xylene) with an unknown concentration, and the root mean square error of calibration was observed to be 1.90, as shown in Figure 5a. In the PLSR analysis, the values of the determination coefficient (R^2) were observed to be 0.99 and 0.99 for the calibration and prediction model, respectively.

2.5. PREDICTION OF UNKNOWN CONCENTRATIONS

The PLSR was applied on the unknown/blind sample (M5), and it predicted the unknown concentration to be 51.25% (v/v), as shown by the red mark in Figure 5b. Table 2 provides

Table 2. Information on how well the PLSR Model Performed while Estimating the Concentration of a few Blind or Unknown Samples that were Chosen at Random

| sample name | calculated concentration | predicted concentration | RMSEC | RMSEP |
|-------------|--------------------------|-------------------------|-------|-------|
| M2 | 20% | 22.13 | 2.31 | 2.78 |
| M5 | 50% | 51.25% | 1.90 | 3.87 |
| M7 | 70% | 70.57 | 1.56 | 2.23 |
| M9 | 90% | 92.57 | 2.45 | 3.45 |

the PLSR estimation of a few randomly chosen unknown/blind samples. When applying the PLSR model, the Raman spectra are not normalized to preserve the link between intensity and concentration.³⁰ Each Raman spectrum is used individually and presented in the predictive model to estimate the unknown concentration; therefore, the prediction results are shown as the mean standard deviation.

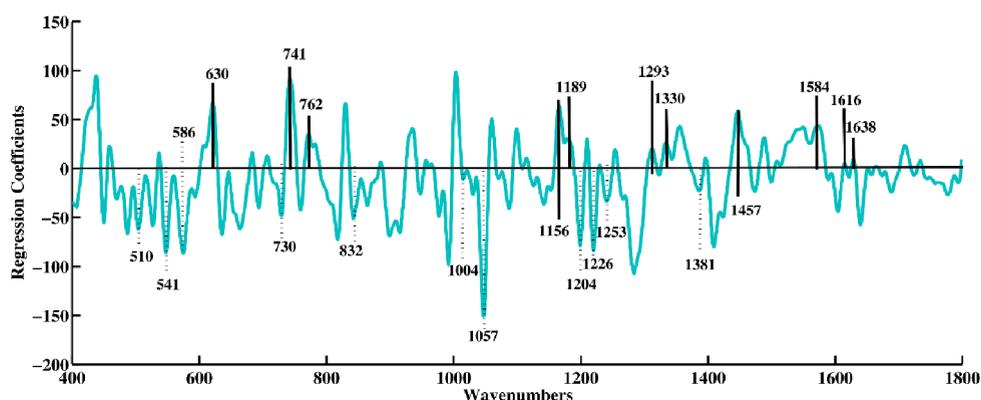


Figure 6. Regression coefficients for the Raman spectral data of all samples of commercial methyl eugenol along with the adulterant (xylene).

2.6. PLS REGRESSION COEFFICIENT

Figure 6 displays the regression coefficients obtained using the PLSR model for the Raman spectral data of all samples. The positive features are observed at 630 (ring deformation), 741 (ring deformation), 762 (deformation of ring), 1156 (C–C bond stretching), 1189 (ring breathing), 1293 (CH₂ twisting vibration), 1330 (twisting vibration of CH₂), 1457 (deformation of CH₂ and CH₃), 1584 (stretching vibration of aromatic C=C bond), 1616 (stretching vibration in conjugated C=C bond), and 1638 cm⁻¹ (C=CH wagging vibration). These features are associated with the Raman spectral characteristic peaks of methyl eugenol. However, the negative features observed here include 510 (ortho-xylene: out of plane bending of C–C–C bond), 541 (ring breathing of meta xylene), 586 (ortho-xylene: ring breathing), 730 (out of plane bending of C–C–C bond in meta xylene ring), 832 (para-xylene: ring bending and out-of-plane bending of C–H bond), 1004 (ring bending of m-xylene), 1057 (ring breathing of ortho xylene), 1204 (meta-xylene: ring bending and in-plane bending of CH), 1226 (CH₃ bond stretching + in-plane bending of CH bond and ring bending of ortho xylene), 1253 (stretching vibrations of the C=C bond and the ring bending of meta xylene), and 1381 cm⁻¹ (C–H bond stretching in meta xylene ring). These features are associated with Raman spectral characteristic peaks of xylene, as their intensities gradually increase with the increase in the concentration of xylene in different samples. Notably, the Raman spectral features found in the regression coefficient confirm the spectral characteristic peaks that appeared in the mean Raman spectra shown in Figure 6.

“Partial least squares (PLS) regression coefficients and Principal component analysis (PCA) loadings are both numerical values used in chemometric data analysis for multivariate analysis of spectroscopic data, including Raman spectra. They are related but serve different purposes, and that’s why they may point to different peaks or features in the data.”⁴²

PLS is a supervised regression technique used for building predictive models when there is a relationship between predictor variables (spectral data) and a response variable (e.g., concentration of a compound of interest). PLS regression coefficients show the importance or contribution of each spectral data point (Raman shift in this case) to the prediction of the response variable.⁴³

The values in the PLS regression coefficients indicate the direction and strength of the relationship between the spectral data and the response variable. Positive values mean an

increase in the response with an increase in spectral intensity, and negative values mean a decrease.⁴⁴

PCA is an unsupervised technique used for dimensionality reduction and exploration of the underlying structure in the data. PCA loadings indicate how each spectral data point contributes to the variance in the data. They help identify patterns and groupings within the data. PCA loadings are not explicitly focused on predicting a specific response variable; instead, they help to understand the overall data structure.⁴⁵

The reason they may point to different peaks or features in Raman data is because they serve different purposes. In some cases, there may be a correspondence between the two, especially if the major sources of variance in data align with the spectral regions that are most important for predicting the response variable. However, the differences in their emphasis and objectives make a direct comparison between PLS and PCA loadings challenging.

Both PLS and PCA can be used in the analysis. First, PCA can be used for data exploration and dimensionality reduction to identify general patterns. Then, PLS can be applied to establish a predictive model for your specific response variable.”

3. CONCLUSIONS

This work demonstrates the capability of Raman spectroscopy along with multivariate statistical methods such as principal component analysis (PCA) and partial least squares regression (PLSR) for the identification and quantification of commercial methyl eugenol along with xylene (an adulterant) in all samples. The characteristic Raman spectral features of commercial methyl eugenol along with xylene (an adulterant) are successfully identified in the mean Raman spectra. PCA is found to be useful for differentiating Raman spectral data sets of different concentrations of methyl eugenol along with xylene in prepared formulations. Moreover, the PLSR model is used to analyze the quantitative relationship between different concentrations of methyl eugenol along with xylene and their Raman spectral features. The PLSR model has been performed to predict the concentration of an unknown sample, with a reliability (R^2) of 0.99 and root-mean-square errors of calibration (RMSEC) and prediction (RMSEP) found to be 1.90 and 3.86, respectively. These results demonstrate that Raman spectroscopy can be employed for the efficient and accurate quantitative detection of adulterants in pesticide products.

4. MATERIALS AND METHODS

4.1. SAMPLE PREPARATION

Commercial methyl eugenol was obtained from a local market in Faisalabad, Pakistan. Xylene (the sum of ortho, meta, and para xylene) (98.5%) was procured from Daejung Chemicals and Metals Co., Ltd., Korea and was used as such without any other purification. Samples of commercial methyl eugenol adulterated with xylene were prepared by adding xylene with different concentrations, as shown in Table 3.

Table 3. Details of Different Volumes of Xylene and Commercial Methyl Eugenol for the Preparation of Different Samples

| sample name | volume of methyl eugenol | volume of xylene | total volume (μL) | total percentage (%) |
|-------------------|--------------------------|-------------------|--------------------------------|----------------------|
| [M ₀] | 0 μL | 100 μL | 100 | 0 |
| [M1] | 10 μL | 90 μL | 100 | 10 |
| [M2] | 20 μL | 80 μL | 100 | 20 |
| [M3] | 30 μL | 70 μL | 100 | 30 |
| [M4] | 40 μL | 60 μL | 100 | 40 |
| [M5] | 50 μL | 50 μL | 100 | 50 |
| [M6] | 60 μL | 40 μL | 100 | 60 |
| [M7] | 70 μL | 30 μL | 100 | 70 |
| [M8] | 80 μL | 20 μL | 100 | 80 |
| [M9] | 90 μL | 10 μL | 100 | 90 |
| [M10] | 100 μL | 0 μL | 100 | 100 |

4.2. RAMAN SPECTRAL DATA ACQUISITION

The Raman spectral acquisition of different concentrations of commercial methyl eugenol along with adulterant (M0–M10) was performed in the liquid form. All spectral measurements were carried out using an Optosky China Raman micro spectrometer (ATR8300BS), with a 785 nm diode laser source. For this purpose, the aluminum slide was used to place the sample, where each sample (about 50 μL) was loaded by using the micropipette in the small groove on an aluminum slide in order to get the Raman spectra. A 40 \times objective lens was utilized to transfer the diode laser to the sample with a 50 mW laser power. This instrument used an air-cooled charge-coupled detector (CCD) to decrease the electrical noise. To ensure the consistency, reliability, and robustness of the data, the same conditions were provided for the spectral acquisition of each sample. Fifteen spectra of each sample were recorded in the 200–1800 cm^{-1} wavenumber range, with an integration duration of 30 s, to acquire each spectrum.

4.3. RAMAN SPECTRAL DATA PREPROCESSING

For further preprocessing and data analysis of all samples, the raw spectral data were imported into MATLAB 7.8.0 R2009a (The MathWorks, USA) and preprocessed using in house developed algorithms.³¹ The Raman spectrum data of various samples (M0–M10) were collected in a matrix and preprocessed collectively to avoid any bias in data preprocessing.⁴⁶ This spectral data preprocessing includes smoothing (Savitzky-Golay filtering), baseline correction, substrate removal, and vector normalization.³¹ Furthermore, the rubber band algorithm for baseline correction was applied to all spectra, and the substrate (aluminum spectrum) was removed from the spectral data of all samples.³²

4.4. RAMAN SPECTRAL DATA ANALYSIS

Multivariate statistical analysis approaches such as principal component analysis (PCA) and partial least-squares regression analysis (PLSR) were used to qualitatively and quantitatively analyze the Raman spectral data of all the samples. The PC loading can be considered as orthogonal variability dimensions that help to separate the different spectral data sets along their coefficients because each spectrum of adulterated methyl eugenol scores along these dimensions.

Principal component analysis is a widely used technique for reducing the dimensionality of data while maintaining the variance among the data set.⁴⁷ In the context of the present study, PCA was applied to the Raman spectral data to identify patterns and reduce the data to a smaller set of components (principal components or PCs) that explain the majority of the variance in the data set. The first PC (PC1) captures the most variance, the second PC (PC2) captures the second most, and so on.

General mathematical representation of PCA is as follows:

$$tk(i) = x(i) \cdot w(k)$$

PC1 is a linear combination of the original features (variables) of the data set. A data set with n features, PC1 is given by

$$PC1 = w11^*X1 + w12^*X2 + \dots + w1n^*Xn$$

where PC1 is the value of the first principal component. $X1$, $X2$, ..., and Xn are the original features. $w11$, $w12$, ..., and $w1n$ are the weights (loadings) assigned to each original feature in PC1. These weights are chosen such that PC1 explains the maximum variance in the data.

PC2 is also a linear combination of the original features and is orthogonal (uncorrelated) to PC1. If you have n features, PC2 is given by

$$PC2 = w21^*X1 + w22^*X2 + \dots + w2n^*Xn$$

where PC2 is the value of the second principal component. $X1$, $X2$, ..., Xn are the original features. $w21$, $w22$, ..., and $w2n$ are the weights assigned to each original feature in PC2. These weights are chosen such that PC2 explains the second most variance in the data and is orthogonal to PC1. The values of $w11$, $w12$, ..., $w1n$ and $w21$, $w22$, ..., $w2n$ are determined through the eigenvectors of the covariance matrix of the original data. The exact values of these weights will depend on the specific data and the implementation of PCA used as they are calculated during the PCA process.

Partial least-squares regression is a multivariate technique used to establish a predictive model when there is a relationship between predictor variables (spectral data) and a response variable (e.g., concentration). PLSR combines aspects of principal component analysis and multiple linear regression.

General mathematical representation for the PLSR is as follows:

In its simplest form, a linear model specifies the (linear) relationship between a dependent (response) variable Y , and a set of predictor variables, the X 's, so that.

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_pX_p$$

Each independent variable (Raman spectral data) is regressed against the relevant concentrations of various samples of methyl eugenol and xylene (adulterant), which

are labeled as dependent factors, according to the regression parameters.⁴⁸

The fitness of the prediction model was assessed by using the root mean square error of cross-validation (RMSECV). Leave one sample (all Raman spectra of one sample), out cross validation (LOOCV) was used to obtain the least number of latent variables and to maximize statistical relevance without overfitting the spectral data. A LOOCV technique has been used to ensure that all Raman spectra recorded from a sample are either contained in the calibration or the test data sets but must not be included in both, to prevent bias during the analysis.

AUTHOR INFORMATION

Corresponding Authors

Muhammad Irfan Majeed – Department of Chemistry, University of Agriculture Faisalabad, Faisalabad 38000, Pakistan; orcid.org/0000-0003-0506-6060; Email: irfan.majeed@uaf.edu.pk

Najah Alwadie – Department of Physics, College of Science, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia; Email: NHAlwadie@pnu.edu.sa

Haq Nawaz – Department of Chemistry, University of Agriculture Faisalabad, Faisalabad 38000, Pakistan; orcid.org/0000-0003-2739-4735; Email: haqchemist@yahoo.com

Authors

Muntaha Anwar – Department of Chemistry, University of Agriculture Faisalabad, Faisalabad 38000, Pakistan

Gull Rimsha – Department of Chemistry, University of Agriculture Faisalabad, Faisalabad 38000, Pakistan

Muhammad Zeeshan Majeed – Department of Entomology, College of Agriculture, University of Sargodha, Sargodha 40100, Pakistan

Nosheen Rashid – Department of Chemistry, University of Education, Faisalabad Campus, Faisalabad 38000, Pakistan

Fareeha Zafar – Department of Chemistry, University of Agriculture Faisalabad, Faisalabad 38000, Pakistan

Ali Kamran – Department of Chemistry, University of Agriculture Faisalabad, Faisalabad 38000, Pakistan

Muhammad Wasim – Department of Chemistry, University of Agriculture Faisalabad, Faisalabad 38000, Pakistan

Nasir Mehmood – Department of Chemistry, University of Agriculture Faisalabad, Faisalabad 38000, Pakistan

Ifra Shabbir – Department of Chemistry, University of Agriculture Faisalabad, Faisalabad 38000, Pakistan

Muhammad Imran – Department of Chemistry, Faculty of Science, King Khalid University, Abha 61413, Saudi Arabia

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acsomega.3c06335>

Author Contributions

*M.A. and G.R. have contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2024R478), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. M. Imran expresses his

appreciation to the Deanship of Scientific Research at King Khalid University Saudi Arabia for funding through the research groups program under grant number R.G.P. 2/522/44.

REFERENCES

- (1) Xu, H.-X.; Zheng, X.-S.; Yang, Y.-J.; Tian, J.-C.; Lu, Y.-H.; Tan, K.-H.; Heong, K.-L.; Lu, Z.-X. Methyl eugenol bioactivities as a new potential botanical insecticide against major insect pests and their natural enemies on rice (*art*). *Crop Prot.* **2015**, *72*, 144–149.
- (2) Maurya, S.; Chandra, M.; Yadav, R. K.; Narnoliya, L. K.; Sangwan, R. S.; Bansal, S.; Sandhu, P.; Singh, U.; Kumar, D.; Sangwan, N. S. Interspecies comparative features of trichomes in *Ocimum* reveal insights for biosynthesis of specialized essential oil metabolites. *Protoplasma* **2019**, *256* (4), 893–907.
- (3) Tan, K. H.; Nishida, R. Methyl eugenol: its occurrence, distribution, and role in nature, especially in relation to insect behavior and pollination. *J. Insect Sci.* **2012**, *12*, 1.
- (4) Kaul, S.; Wani, M.; Dhar, K. L.; Dhar, M. K. Production and GC-MS trace analysis of methyl eugenol from endophytic isolate of *Alternaria* from rose. *Ann. Microbiol.* **2008**, *58* (3), 443–445.
- (5) Vargas, R. I.; Miller, N. W.; Stark, J. D. Field trials of spinosad as a replacement for naled, DDVP, and malathion in methyl eugenol and cue-lure bucket traps to attract and kill male oriental fruit flies and melon flies (Diptera: Tephritidae) in Hawaii. *J. Econ. Entomol.* **2003**, *96* (6), 1780–1785.
- (6) Lahlou, S.; Figueiredo, A. F.; Magalhães, P. J. C.; Leal-Cardoso, J. H.; Gloria, P. D. Cardiovascular effects of methyleugenol, a natural constituent of many plant essential oils, in normotensive rats. *Life Sci.* **2004**, *74* (19), 2401–2412.
- (7) Tripathi, A. K.; Mishra, S. Plant monoterpenoids (prospective pesticides). In *Ecofriendly Pest Management for Food Security*; Academic Press, 2016; pp 507–524. DOI: .
- (8) Waykole, P.; Badekar, R.; Lokhande, R.; Nemade, H. Structural Studies of Novel Synthesized Compounds from Methyleugenol with Various Acid Derivatives of Indole. *Adv. Innov. Res.* **2018**, *5*, 25–28.
- (9) Bendre, R. S.; Rajput, J. D.; Bagul, S. D.; Karandikar, P. Outlooks on medicinal properties of eugenol and its synthetic derivatives. *Nat. Prod. Chem. Res.* **2016**, *4* (3), 1–6.
- (10) Goswami, P.; Verma, S. K.; Chauhan, A.; Venkatesha, K. T.; Verma, R. S.; Singh, V. R.; Darokar, M. P.; Chanotiya, C. S.; Padalia, R. C. Chemical composition and antibacterial activity of *Melaleuca bracteata* essential oil from India: A natural source of methyl eugenol. *Nat. Prod. Commun.* **2017**, *12* (6), 1934578X1701200633.
- (11) Faisal, M.; Prasad, L.; Icar-Iisr, R. A potential source of methyl-eugenol from secondary metabolite of *Rhizopus oryzae* 6975. *Int. J. Appl. Biol. Pharm. Technol.* **2016**, *7*, 187–192.
- (12) Purkait, A.; Hazra, D. K. Biodiesel as a carrier for pesticide formulations: A green chemistry approach. *Int. J. Pest Manage.* **2020**, *66* (4), 341–350.
- (13) Cox, C.; Sorgan, M. Unidentified inert ingredients in pesticides: Implications for human and environmental health. *Environ. Health Perspect.* **2006**, *114* (12), 1803–1806.
- (14) Beggel, S.; Werner, I.; Connon, R. E.; Geist, J. P. Sublethal toxicity of commercial insecticide formulations and their active ingredients to larval fathead minnow (*Pimephales promelas*). *Sci. Total Environ.* **2010**, *408* (16), 3169–3175.
- (15) Emara, A. M.; Draz, E. I. Immunotoxicological study of one of the most common over-the-counter pyrethroid insecticide products in Egypt. *Inhalation Toxicol.* **2007**, *19* (12), 997–1009.
- (16) Choudhary, A.; Gupta, N.; Hameed, F.; Choton, S. An overview of food adulteration: Concept, sources, impact, challenges and detection. *Int. J. Chem. Stud.* **2020**, *8* (1), 2564–2573.
- (17) Bayoumi, A. E. Counterfeit pesticides. *ACS Chem. Health Saf.* **2021**, *28* (4), 232–237.
- (18) Zikankuba, V. L.; Mwanyika, G.; Ntwenya, J. E.; James, A. Pesticide regulations and their malpractice implications on food and environment safety. *Cogent Food Agric.* **2019**, *5* (1), 1601544.

- (19) Hussain, S.; Khalid, M. A.; Ahmad, M. J.; ul Haq, E.; Nawaz, S.; Yaseen, G.; Riaz, S.; Arif, M. ASSESSMENT OF PESTICIDES FOR QUALITY CONTROL UNDER PESTICIDE REGULARITY PROGRAMME. *J. Agric. Res.* **2020**, *58* (3), 171–176.
- (20) Miszczyk, M.; Płonka, M.; Bober, K.; Dołowy, M.; Pyka, A.; Pszczolińska, K. Application of chemometric analysis based on physicochemical and chromatographic data for the differentiation origin of plant protection products containing chlorpyrifos. *J. Environ. Sci. Health, Part B* **2015**, *50* (10), 744–751.
- (21) Su-Yan, Z.; Yu, G.; Yin-Long, G.; Hao, W.; Long, L. Identification of Novel Pesticides and Impurities by the Combination of LC-MS with GC-MS Analysis. *Chin. J. Chem.* **2005**, *23* (7), 870–874.
- (22) Siebers, J.; Besinger-Riedel, A.; Vinke, C. Determination of active substances, co-formulants and impurities in plant protection products using high performance liquid chromatography and gas chromatography. *Journal für Verbraucherschutz und Lebensmittelsicherheit* **2014**, *9* (2), 137–144.
- (23) Li, Z.; Li, Y.; Zhang, X.; Zhou, J.; Shen, S. Identification, isolation, characterization, and synthesis of impurities of pesticide spirotetramat. *J. Environ. Sci. Health, Part B* **2022**, *57* (9), 765–773.
- (24) Balayiannis, G. P.; Anagnostopoulos, H.; Kellidou, I. Facile and Rapid Determination of Contamination in Sulphur Pesticide Formulations by Liquid Chromatography–Tandem Mass Spectrometry. *Bull. Environ. Contam. Toxicol.* **2009**, *82* (2), 133–136.
- (25) Marczewska, P.; Płonka, M.; Rolnik, J.; Sajewicz, M. Determination of azoxystrobin and its impurity in pesticide formulations by liquid chromatography. *J. Environ. Sci. Health, Part B* **2020**, *55* (7), 599–603.
- (26) Khattab, G. M.; Abdelghany, W.; Abdelmegeed, M. I.; Attalah, I. Monitoring of counterfeit abamectin pesticide products in Egypt. *Arab Universities Journal of Agricultural Sciences* **2020**, *28* (2), 601–615.
- (27) Gallart-Mateu, D.; Armenta, S.; de la Guardia, M. Implementing the contamination prevention programs in the pesticide industry by infrared spectroscopy. *Talanta* **2014**, *119*, 312–319.
- (28) Liu, Y.; Liu, T. Determination of pesticide residues on the surface of fruits using micro-Raman spectroscopy. In *International Conference on Computer and Computing Technologies in Agriculture*; Springer, 2010; pp 427–434.
- (29) Bakkar, M. A.; Nawaz, H.; Majeed, M. I.; Naseem, A.; Ditta, A.; Rashid, N.; Ali, S.; Bajwa, J.; Bashir, S.; Ahmad, S.; et al. Raman spectroscopy for the qualitative and quantitative analysis of solid dosage forms of Sitagliptin. *Spectrochim. Acta, Part A* **2021**, *245*, 118900.
- (30) Shafaq, S.; Majeed, M. I.; Nawaz, H.; Rashid, N.; Akram, M.; Yaqoob, N.; Tariq, A.; Shakeel, S.; ul Haq, A.; Saleem, M.; et al. Quantitative analysis of solid dosage forms of Losartan potassium by Raman spectroscopy. *Spectrochim. Acta, Part A* **2022**, *272*, 120996.
- (31) Nawaz, H.; Rashid, N.; Saleem, M.; Asif Hanif, M.; Irfan Majeed, M.; Amin, I.; Iqbal, M.; Rahman, M.; Ibrahim, O.; Baig, S. Prediction of viral loads for diagnosis of Hepatitis C infection in human plasma samples using Raman spectroscopy coupled with partial least squares regression analysis. *J. Raman Spectrosc.* **2017**, *48* (5), 697–704.
- (32) Mahmood, T.; Nawaz, H.; Ditta, A.; Majeed, M.; Hanif, M.; Rashid, N.; Bhatti, H.; Nargis, H.; Saleem, M.; Bonnier, F.; et al. Raman spectral analysis for rapid screening of dengue infection. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2018**, *200*, 136–142.
- (33) Ashraf, M. N.; Majeed, M. I.; Nawaz, H.; Iqbal, M. A.; Iqbal, J.; Iqbal, N.; Hasan, A.; Rashid, N.; Abubakar, M.; Nawaz, M. Z.; et al. Raman spectroscopic characterization of selenium N-heterocyclic carbene compounds. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2022**, *270*, 120823.
- (34) Armenta, S.; Garrigues, S.; de la Guardia, M. Determination of iprodione in agrochemicals by infrared and Raman spectrometry. *Anal. Bioanal. Chem.* **2007**, *387* (8), 2887–2894.
- (35) Takayasu, T.; Ohshima, T.; Kondo, T. Rapid analysis of pesticide components, xylene, o-dichlorobenzene, cresol and dichlorvos, in blood and urine by pulse heating–gas chromatography–mass spectrometry. *Leg. Med.* **2001**, *3* (3), 157–161.
- (36) Qi, Y.; Ma, C.; Wan, M.; Liu, Z.; Zheng, L.; Jin, F.; Jin, M.; She, Y.; Wang, J.; Wang, S. Multiresidue Determination of Six Pesticide Adjuvants in Characteristic Minor Crops Using QuEChERS Method and Gas Chromatography–Mass Spectrometry. *ChemistrySelect* **2019**, *4* (1), 66–70.
- (37) Jinnō, K.; Muramatsu, T.; Saito, Y.; Kiso, Y.; Magdic, S.; Pawliszyn, J. Analysis of pesticides in environmental water samples by solid-phase micro-extraction—high-performance liquid chromatography. *J. Chromatogr. A* **1996**, *754* (1–2), 137–144.
- (38) Vedad, J.; Reilly, L.; Desamero, R. Z. B.; Mojica, E.-R. E. Quantitative Analysis of Xylene Mixtures Using a Handheld Raman Spectrometer. In *Raman Spectroscopy in the Undergraduate Curriculum*; ACS Publications, 2018; pp 129–152.
- (39) Wang, L.-H.; Chen, J.-X.; Wang, C.-C. Rapid quantitative analysis of suspected fragrance allergens in between commercial essential oils and using attenuated total reflectance–infrared (ATR–IR) spectroscopy. *J. Essent. Oil Res.* **2014**, *26* (3), 185–196.
- (40) Dhakal, S.; Li, Y.; Peng, Y.; Chao, K.; Qin, J.; Guo, L. Prototype instrument development for non-destructive detection of pesticide residue in apple surface using Raman technology. *J. Food Eng.* **2014**, *123*, 94–103.
- (41) Boaroto, J. A.; Fernandes, A. B.; Moreira, L. M.; Silveira Jr, L.; José de Lima, C. Identification and quantification of β -caryophyllene in copaiba oil using Raman spectroscopy. *Instrum. Sci. Technol.* **2018**, *46* (3), 265–276.
- (42) Dunn, W., III; Scott, D.; Glen, W. G. Principal components analysis and partial least squares regression. *Tetrahedron Comput. Methodol* **1989**, *2* (6), 349–376.
- (43) Mehmood, T.; Liland, K. H.; Snipen, L.; Sæbø, S. A review of variable selection methods in partial least squares regression. *Chemom. Intell. Lab. Syst.* **2012**, *118*, 62–69.
- (44) Jiang, J.-H.; Berry, R. J.; Siesler, H. W.; Ozaki, Y. Wavelength interval selection in multicomponent spectral analysis by moving window partial least-squares regression with applications to mid-infrared and near-infrared spectroscopic data. *Anal. Chem.* **2002**, *74* (14), 3555–3565.
- (45) Hasan, B. M. S.; Abdulzееz, A. M. A review of principal component analysis algorithm for dimensionality reduction. *Journal of Soft Computing and Data Mining* **2021**, *2* (1), 20–30.
- (46) Nasir, S.; Majeed, M. I.; Nawaz, H.; Rashid, N.; Ali, S.; Farooq, S.; Kashif, M.; Rafiq, S.; Bano, S.; Ashraf, M. N.; et al. Surface enhanced Raman spectroscopy of RNA samples extracted from blood of hepatitis C patients for quantification of viral loads. *Photodiagn. Photodyn. Ther.* **2021**, *33*, 102152.
- (47) Huda, N. U.; Wasim, M.; Nawaz, H.; Majeed, M. I.; Rashid, N.; Javed, M. R.; Iqbal, M. A.; Tariq, A.; Hassan, A.; Akram, M. W. Synthesis, Characterization, and Evaluation of Antifungal Activity of 1-Butyl-3-hexyl-1H-imidazol-2(3H)-selenone by Surface-Enhanced Raman Spectroscopy. *ACS Omega* **2023**, *8* (39), 36460–36470.
- (48) Makki, A. A.; Bonnier, F.; Respaud, R.; Chtara, F.; Tfayli, A.; Tauber, C.; Bertrand, D.; Byrne, H. J.; Mohammed, E.; Chourpa, I. Qualitative and quantitative analysis of therapeutic solutions using Raman and infrared spectroscopy. *Spectrochim. Acta, Part A* **2019**, *218*, 97–108.