

An Integrated Authentication Analysis of *Citrus aurantium* L. Essential Oil Based on FTIR Spectroscopy and Chemometrics with Tuning Parameters

Florentinus Dika Octa Riswanto^{1*}, Anjar Windarsih², Dina Christin Ayuning Putri³ and Michael Raharja Gani¹

1. Division of Pharmaceutical Analysis and Medicinal Chemistry, Faculty of Pharmacy, Campus III Paingan, Universitas Sanata Dharma, Maguwoharjo, Sleman, Yogyakarta 55282, Indonesia
2. Research Center for Food Technology and Processing (PRTPP), National Research and Innovation Agency (BRIN), Yogyakarta 55861
3. Division of Pharmaceutical Technology, Faculty of Pharmacy, Campus III Paingan, Universitas Sanata Dharma, Maguwoharjo, Sleman, Yogyakarta 55282, Indonesia

Article Info

Submitted: 29-06-2022

Revised: 21-10-2022

Accepted: 30-03-2023

*Corresponding author
Florentinus Dika Octa R.

Email:
dikaocta@usd.ac.id

ABSTRACT

Citrus aurantium L. essential oil or Orange Oil (OO) has gained popularity recently due to its benefit for human health. An “economically motivated adulteration” can potentially occur to achieve more profit in the market. On the other hand, cheaper oil, such as Coconut Oil (CO), is commonly used as an adulterant. This study aims to perform an authentication analysis of OO by FTIR spectroscopy and chemometrics. The exploratory data analysis applied the principal component analysis at the initial stage of authentication analysis. Multivariate calibration of Principal Component Regression (PCR) and Partial Least Squares Regression (PLSR) was constructed from five pre-processed FTIR spectral data. The PCR model using Standard Normal Variate (SNV) spectra was selected as the best prediction model for OO, whereas the PLSR model using SNV spectra was sorted for the best prediction model for CO. SNV spectra of OO, CO, and binary mixture of OO+CO was used to generate sparse partial least squares-discriminant analysis (SPLS-DA) model. Component number three with “keepX” for components 1, 2, and 3 were 1, 5, and 1, respectively, were selected along with the maximum distance approach to construct the discriminant model. The final sPLS-DA model explained the total variances of 94% with satisfaction separability of 100%, 97.8%, and 100% for OO, CO, and OO+CO, respectively. In conclusion, FTIR spectroscopy and chemometrics with tuning parameters can authenticate *Citrus aurantium* L. essential oil.

Keywords: authentication, chemometrics, FTIR spectroscopy, *Citrus aurantium* L.

INTRODUCTION

Essential oils, a group of volatile and aromatic oils from natural sources, can be obtained by steam distillation process due to their volatility (Hyldgaard *et al.*, 2012). Each type of essential oil is named based on its sources, the extraction method, and the odor produced (Gautam *et al.*, 2021). The essential oils have been used not only for culinary purposes but also for different purposes such as beauty care products (Naeem *et al.*, 2018), pharmaceutical and therapeutic products (Edris, 2007), pest repellants (Lee, 2018), anti-bacterial

agents (Wińska *et al.*, 2019), and food packaging materials (Refaie *et al.*, 2020).

Ancient people in India, China, and Egypt (Krishna *et al.*, 2000; Manniche, 1999) have known aromatherapy as a complementary alternative since approximately 6,000 years ago and utilized this oil. Essential oil aromatherapies have become popular in recent years because this oil can be used as a therapeutic agent to treat various health problems (Ali *et al.*, 2015). The marketed aromatherapies usually contain essential oils in high concentrations extracted from flowers, leaves,

stems, fruits, and roots (Dunning, 2013). A study subject named aroma science therapy has gained importance. It is due to its high popularity, great demand, and awareness to study several aromatherapy aspects (Esposito *et al.*, 2014). *Citrus aurantium* L., commonly known as bitter orange or sour orange, is one species in the *Rutaceae* family and genus *Citrus* that can be easily found and cultivated in Italy, Spain, and the United States (Maksoud *et al.*, 2021). As reported in the previous study, Citrus plants are associated with herbal medicines in Asian countries such as China, Korea, and Japan (Lv *et al.*, 2015). *Citrus aurantium* L. essential oil was reported to have pharmacological activity and various chemical compositions due to the growing area and seasonal variation but mostly contained marker compounds, namely limonene, linalool, and β -myrcene (Boussaada & Chemli, 2007; Suntar *et al.*, 2018). Chemical compositions of *Citrus aurantium* L. were related to several activities, such as larvicidal activity (Sanei-Dehkordi *et al.*, 2016), antimicrobial activity (Bnina *et al.*, 2019), and fumigant activity (Djebbi *et al.*, 2021). In a recent study, microencapsulation of *Citrus aurantium* L. essential oil from orange fruit peel was successfully prepared, characterized, and proved as an antimicrobial and antioxidant agent (de Araújo *et al.*, 2020).

The increased interest in using *Citrus aurantium* L. essential oil leads to the potential “economically motivated adulteration” to gain more profit in the market (Johnson, 2014). *Citrus aurantium* L. essential oil is 7-10 times more expensive than other natural oils, such as coconut and palm oil. As a result, *Citrus aurantium* L. essential oil is easily substituted with other cheaper oils (Rohman *et al.*, 2014). Hence, the detection and quantification of oil adulterants are significant to develop.

Vibrational spectroscopy techniques such as Fourier Transform Infrared (FTIR) spectroscopy can be employed as a green tool in multicomponent analysis (Moros *et al.*, 2010). Analyzing fish and vegetable oil for qualitative and quantitative aspects applies differentiation and authentication using FTIR spectroscopy (Putri *et al.*, 2019; Rohman & Che Man, 2010, 2011). The exploitation of the employment of the FTIR spectroscopy method occurs because it is fast, sensitive, simple, and non-destructive to sample preparation (Reid *et al.*, 2006). The FTIR spectroscopic method and chemometric techniques are used for the authentication analysis (Irnawati *et al.*, 2021;

Sim & Jeffrey Kimura, 2019). Chemometrics techniques such as Principal Component Analysis (PCA), Principal Component Regression (PCR), Partial Least Squares Regression (PLSR), and Partial Least Squares-Discriminant Analysis (PLS-DA) can be combined with spectroscopy method for authentication purposes (Dzulfiyanto *et al.*, 2018; Hemmateenejad *et al.*, 2007; Nurani *et al.*, 2021; Riswanto *et al.*, 2021, 2022). This study aimed to perform an authentication analysis of *Citrus aurantium* L. essential oil or Orange Oil (OO) because the publications reporting the employment of FTIR spectroscopy and chemometrics along with tuning parameters are limited. PCA explains the exploratory data analysis of OO. The establishment of multivariate calibration of PCR and PLSR is to quantitatively predict the content of OO adulterated with **Coconut Oil (CO)**. Sparse PLS-DA (SPLS-DA) was generated and accompanied by tuning parameters to discriminate pure OO in adulterated samples.

MATERIAL AND METHODS

Pure *Citrus aurantium* L. essential oil, purchased from a local distributor in Yogyakarta, Indonesia, was first produced and distributed from Utah, USA. Meanwhile, a local market in Sleman, Indonesia, provides coconut oil. The solvent used in this study is ethanol obtained from Merckmillipore.

Instrumentation and Software

This study applied a set of Bruker VERTEX 80 FT-IR spectrophotometer and Socorex[®] micropipettes of 20-200 μ L, 100-1000 μ L, and 500-5000 μ L. The data of FTIR spectra were exported into Excel 2016 (Microsoft Inc., USA) and saved as .csv files. PCA used R statistical software version 4.1.1 with ‘factoextra’ and ‘FactoMineR’ packages (Irnawati *et al.*, 2021). Multivariate calibrations and spectral preprocessing applied the ‘pls’ and ‘prospectr’ packages, respectively (Mevik & Wehrens, 2019; Stevens & Ramirez Lopez, 2020). SPLS-DA modeling and tuning were generated from the ‘mixOmics’ package (Rohart *et al.*, 2017).

Calibration and Validation Sample Preparation

A set of calibration and validation solutions containing OO and CO were prepared from both pure oils to obtain 66 concentration variations of calibration solutions and 30 concentration variations of validation solutions. All these solutions were prepared by mixing OO and CO with a concentration range of 0-100% (v/v).

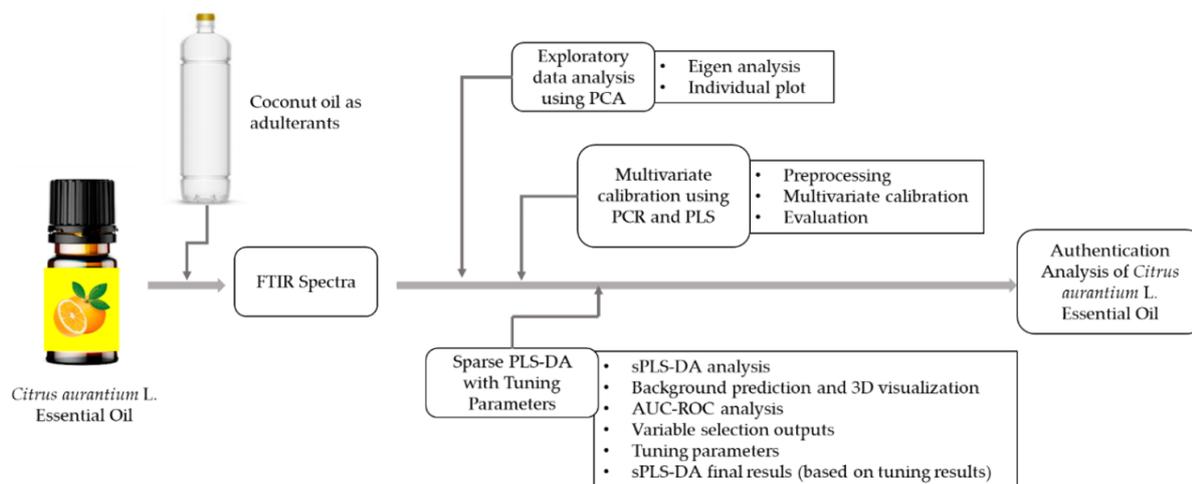


Figure 1. Scheme of authentication analysis of *Citrus aurantium* L. essential oil

Sample Preparation for Discrimination

Solutions for the SPLS-DA study were available in three different classes, namely Orange Oil (OO), Coconut Oil (CO), and a binary mixture containing OO and CO (OO+CO) with the concentration range of 0-100% (v/v). Seven pure OO solutions, seven pure CO solutions, and thirty-two OO+CO solutions were obtained and labeled for constructing the discrimination analysis model.

FTIR Spectra Acquisition

FTIR spectra acquisition uses an FTIR spectrophotometer (Bruker Vertex-80, Germany) equipped with attenuated total reflectance (ATR) as the sampling technique and deuterated triglycine sulfate (DTGS) detector. A sample of pure CO, pure OO, and adulterated CO was placed on ATR crystal and then measured at mid-infrared region (4000-600 cm^{-1}). The FTIR spectra employed absorbance mode using a resolution of 8 cm^{-1} and several scans 32. Measurement of air spectra applied before each sample measurement as the background spectra. Measuring each sample utilized three replicates. The ATR crystal cleansing used ethanol analytical grade after each sample measurement. The FTIR spectra were processed using OPUS Software version 8.5 (Bruker, Germany).

Exploratory Data Analysis using PCA

In this study, the PCA model utilized the finding of intrinsic structures of the multidimensional data. An evaluation examined the

FTIR spectra obtained from the data acquisition stage. Twelve dominant peaks of the OO, CO, and OO+CO analysis built the PCA model accompanied by the eigen analysis evaluation. The presentation of the Scree plot, variable plot, and individual plot visualized the optimal number of components and the initial variables load for constructing the components and graphs the general grouping of the variables explained from the first and second components, respectively.

Multivariate Calibration Techniques

Absorbance data for each wavenumber of calibration and validation solutions were treated and preprocessed into five FTIR spectra types. They were initial spectra, first derivative, second derivative, Standard Normal Variate (SNV), and Savitzky-Golay smoothing (SG) with eleven points window width and polynomial order of 3.

Multivariate calibration models of PCR and PLSR generated a predictive model for two oils. Multivariate calibration model performance was examined by evaluating statistical parameters, including coefficient of determination for calibration (R_{cal}^2), cross-validation (R_{CV}^2), validation (R_{val}^2), root mean square error of calibration (RMSEC), root mean square error of cross-validation (RMSECV) and root mean square error of prediction (RMSEP). The cross-validation process as internal validation applied a leave-one-out technique. The selected multivariate calibration model for each oil was determined by evaluating the R_{cal}^2 , R_{CV}^2 , R_{val}^2 , RMSEC, RMSECV, and RMSEP.

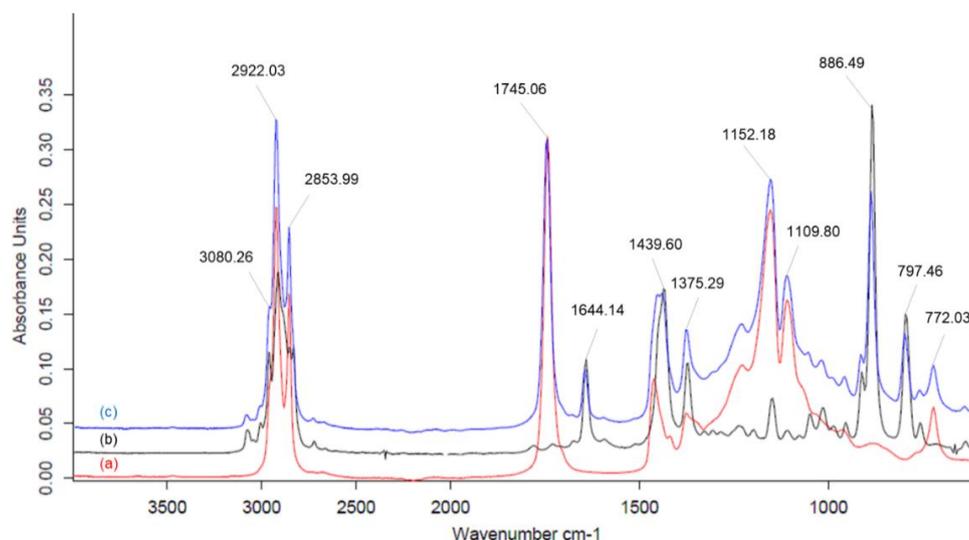


Figure 2. FTIR spectra of coconut oil (a), orange oil (b), and binary mixture of orange oil and coconut oil (c)

SPLS-DA Model Generation with Tuning Parameters

SPLS-DA model utilized the FTIR spectra of OO, CO, and OO+CO. The background prediction and 3D individual plots visualized discrimination models. The model performance evaluation applied the area under the curve - receiver operating characteristics (AUC-ROC). This model optimization was done further by selecting output variables and tuning parameters. The final SPLS-DA results came from considering the classification error rates and feature selection (Figure 1).

RESULTS AND DISCUSSION

This study initially examines the FTIR spectra resulting from the scanning process. This study employs FTIR spectroscopy due to its ability to provide direct information regarding the molecular species present in the oil and functional groups for further chemical identification (Shabanian *et al.*, 2020). The FTIR spectra profiles of OO, CO, and OO+CO (Figure 2). The FTIR spectra applied for functional groups, vibrational modes, and intensities due to their wavenumbers (Table I).

FTIR spectra of OO, CO, and CO+OO are analyzed. 3080, 2922, 2853, 1745, 1644, 1439, 1375, 1152, 1109, 886, 797, and 772 cm^{-1} convey the primary peaks. The peak marker of OO characterizes the vibrational bands at 3080, 1644, 1439, 886, 797, and 772 cm^{-1} . Since OO was composed of terpenes (with (+)-limonene as the significant compound), sesquiterpenes, and aldehydes, it shows that the stretching and bending vibration may include spectral

features arising from $-\text{CH}_3$, $-\text{C}=\text{C}$, and $-\text{C}-\text{H}$ of terpenes (Cebi *et al.*, 2021). These functional groups represented the characteristics of common terpenes, namely α -pinene, β -pinene, champhene, limonene, α -phellandrene, β -phellandrene, and β -myrcene (Derdar *et al.*, 2019; Yang *et al.*, 2017). The manifestation of pyranose skeleton bending presents specified sharp peaks at 772, 733, and 700 cm^{-1} (Simona *et al.*, 2021). On the other hand, the vibrational bands at 2922, 2853, 1745, 1375, 1152, and 1109 cm^{-1} represented the fingerprint pattern of CO as an adulterant. Absorption intensities at 2922, 2853, and 1745 cm^{-1} reveal the presence of ester and aliphatic groups corresponding to triglyceride content in CO (Rohman *et al.*, 2019). Other regions of 1375, 1152, and 1109 cm^{-1} corresponding to the bending and starching vibrations of $-\text{C}-\text{H}$, $-\text{C}-\text{O}$, $-\text{CH}_2-$, and $-\text{CH}_3$ functional groups related to the structure of fatty acids (Amit *et al.*, 2020).

Exploratory Data Analysis using PCA

PCA, a dimensionality reduction technique for large datasets, is commonly applied for exploratory methods in data science (Van Der Maaten *et al.*, 2009). PCA, a pattern recognition algorithm, provides a descriptive tool for data treatment without any distributional assumptions or target attributes (Jolliffe & Cadima, 2016). In analytical chemistry, data output from analytical instruments may be obtained as big data. Recording of absorbance at each wavenumber from oil samples uses an FTIR spectrophotometer in the range of 4000-600 cm^{-1} with a resolution of 8 cm^{-1} .

Table II. The performance of principle component regression (PCR) and partial least squares regression (PLSR) for predicting the content of mixture containing orange oil and coconut oil

| Oils | Multivariate calibration | Type of spectra | Number of components | R _{cal} ² | RMSEC | R _{cv} ² | RMSECV | R _{val} ² | RMSEP |
|-------------|--------------------------|-------------------|----------------------|-------------------------------|--------------|------------------------------|--------------|-------------------------------|--------------|
| Orange oil | PCR | Original | 58 | 1.000 | 0.079 | 0.999 | 0.413 | 0.991 | 1.023 |
| | | First derivative | 64 | 1.000 | 0.013 | 0.998 | 0.444 | 0.991 | 1.026 |
| | | Second derivative | 63 | 1.000 | 0.036 | 0.998 | 0.479 | 0.990 | 1.113 |
| | | SNV | 39 | 0.999 | 0.193 | 0.998 | 0.456 | 0.992 | 0.989 |
| | | SG | 36 | 0.999 | 0.191 | 0.999 | 0.432 | 0.991 | 1.043 |
| | PLSR | Original | 16 | 0.999 | 0.104 | 0.999 | 0.420 | 0.991 | 1.032 |
| | | First derivative | 21 | 1.000 | 0.001 | 0.998 | 0.444 | 0.991 | 1.027 |
| | | Second derivative | 16 | 1.000 | 0.007 | 0.998 | 0.480 | 0.990 | 1.115 |
| | | SNV | 14 | 0.999 | 0.174 | 0.998 | 0.476 | 0.992 | 0.991 |
| | | SG | 25 | 1.000 | 0.011 | 0.999 | 0.433 | 0.992 | 0.956 |
| Coconut oil | PCR | Original | 58 | 1.000 | 0.079 | 0.999 | 0.413 | 0.991 | 1.023 |
| | | First derivative | 64 | 1.000 | 0.014 | 0.998 | 0.444 | 0.991 | 1.026 |
| | | Second derivative | 63 | 1.000 | 0.036 | 0.998 | 0.479 | 0.990 | 1.113 |
| | | SNV | 39 | 0.999 | 0.193 | 0.998 | 0.455 | 0.992 | 0.989 |
| | | SG | 36 | 0.999 | 0.191 | 0.999 | 0.432 | 0.991 | 1.043 |
| | PLSR | Original | 16 | 0.999 | 0.104 | 0.999 | 0.420 | 0.991 | 1.032 |
| | | First derivative | 21 | 1.000 | 0.001 | 0.998 | 0.444 | 0.991 | 1.027 |
| | | Second derivative | 16 | 1.000 | 0.007 | 0.998 | 0.480 | 0.990 | 1.115 |
| | | SNV | 14 | 0.999 | 0.174 | 0.999 | 0.476 | 0.992 | 0.991 |
| | | SG | 25 | 1.000 | 0.011 | 0.999 | 0.433 | 0.992 | 0.956 |

Note: Selected model of calibration for each compound were marked with bold. PCR: Principal Component Regression; PLSR: Partial Least Squares Regression; SNV: Standard Normal Variate; SG: Savitzky-Golay smoothing with polynomial order of 3 and window width of 11 points.

Hence, a data treatment strategy is significant in explaining the intrinsic structures for further evaluation.

The absorbance of twelve primary peaks from the FTIR spectra generates the PCA model. These spectra function as original data without any spectral treatment or preprocessing techniques. Figure 3 presents the screen plot, variable plot, and PCA model individual plot. The screen plot depicts the number of crucial components in generating the PCA model. Meanwhile, the score plot provides information on how the samples relate to each other in two-dimensional visualization. A variable plot, commonly called a loading plot, explains the contribution of each original variable to build the model.

The individual plot successfully separates pure OO from pure CO as adulterants. The adulterated OO+CO is present in various scores or locations because of different concentration preparations. The total explained variances for creating a two-dimensional PCA model is 79.5%, as calculated by Dim1 (61.6%) and Dim2 (17.9%). The term “Dim” or dimension refers to the principal component (PC) as commonly recognized in chemometrics terminology (Kassambara & Mundt, 2017).

Because the FTIR spectral data of twelve primary peaks build the PCA, it has been interesting to evaluate the contribution of each original variable. The variable plot shows a load of each variable, as well as their contribution profiles. The exhibition of the absorbance at 1745, 1645, 1152, 886, and 772 cm⁻¹ is the main contributor to the construction of the PCA model. In conclusion, the vibrational bands of these regions provide specific fingerprint properties for OO and CO. Vibrational bands near 1645, 886, and 772 cm⁻¹ result from the content of OO. Meanwhile, the content of CO contributed to the presence of the bands is near 1745 and 1152 cm⁻¹.

Multivariate Calibrations

In this authentication study, multivariate calibration techniques, namely PCR and PLSR are generated to develop predictive models for quantitative evaluation in OO adulteration. PCR is applied to decrease the predictor's variable number by using their first few principal components selected by cross-validation technique rather than the original variables. Meanwhile, the PLSR employed linear combinations of the predictor variables rather than the original (Kassambara, 2018; Miller *et al.*, 2018).

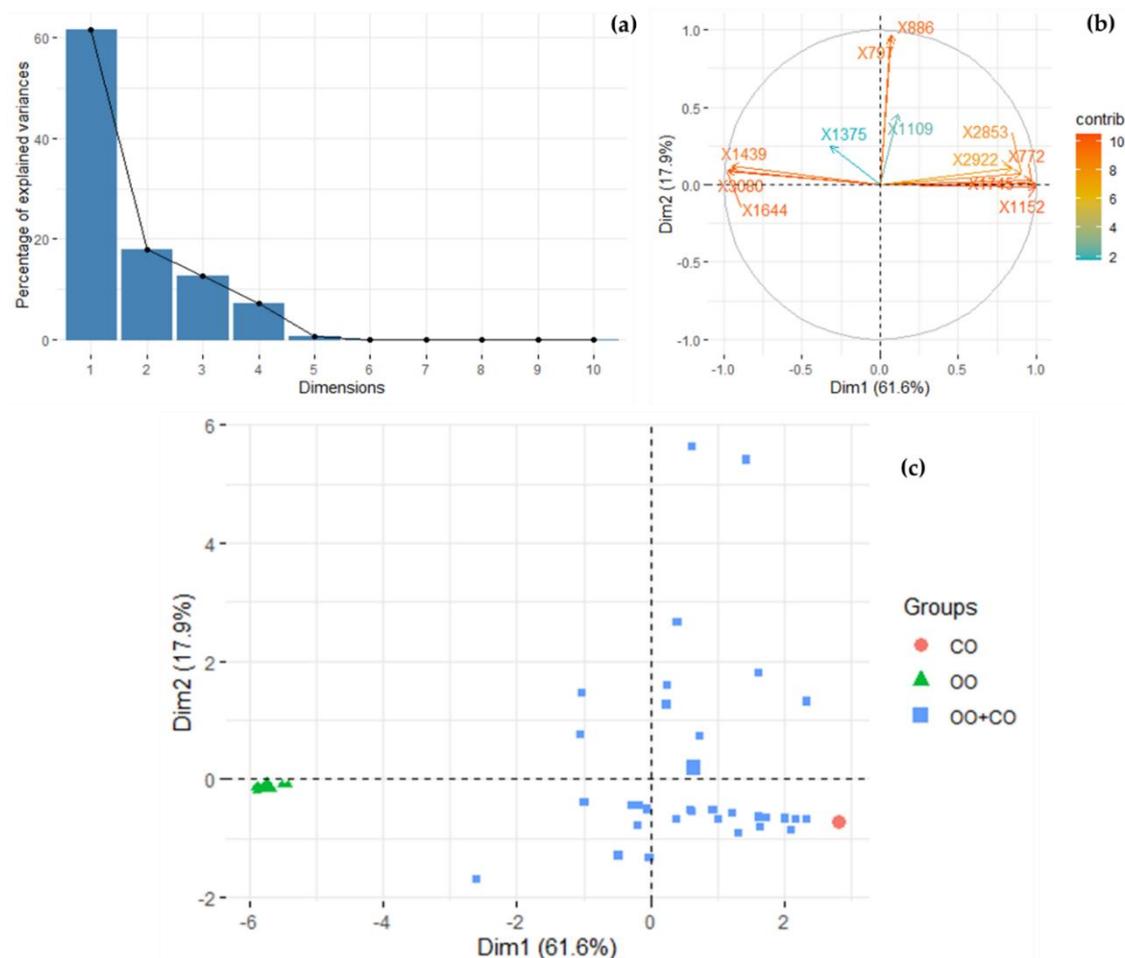


Figure 3. Scree plot (a), variable plot (b), and individual plot (c) of principal component analysis resulted from original spectra at selected wavenumber

FTIR spectral data ranging between 4000-600 nm^{-1} aims to achieve different spectra, including original spectra, first derivative, second derivative, SNV, and SG. The objective of spectral pre-processing is to control the risk of overfitting and boundary complexity. Furthermore, removing undesired variance in the spectra improves the model's predictive ability (Devos *et al.*, 2014).

The PCR model using SNV spectra and the PLSR model using SNV spectra are selected as the best prediction model for OO and CO, respectively. These two models have been considered appropriate due to the highest R^2 values (R_{cal}^2 , R_{CV}^2 , R_{val}^2) and the lowest value of root mean squares error (RMSEC, RMSECV, RMSEP).

In conclusion, the selected models were accurate and precise for calibration and validation models (Rohman *et al.*, 2014).

The predictive model generated from a set of training data refers to the calibration model in chemometrics. Examining R_{cal}^2 and RMSEC values can evaluate this model. The performance of cross-validation as the internal validation of the training data uses leave-one-out techniques. The R_{CV}^2 and RMSECV indicate the quality of the developed model when the internal validation model applies. Since the external validation model constructions come from independent datasets, the predictive ability of the models proves a quantitative determination regarding the value of R_{val}^2 and RMSEP (Figure 4).

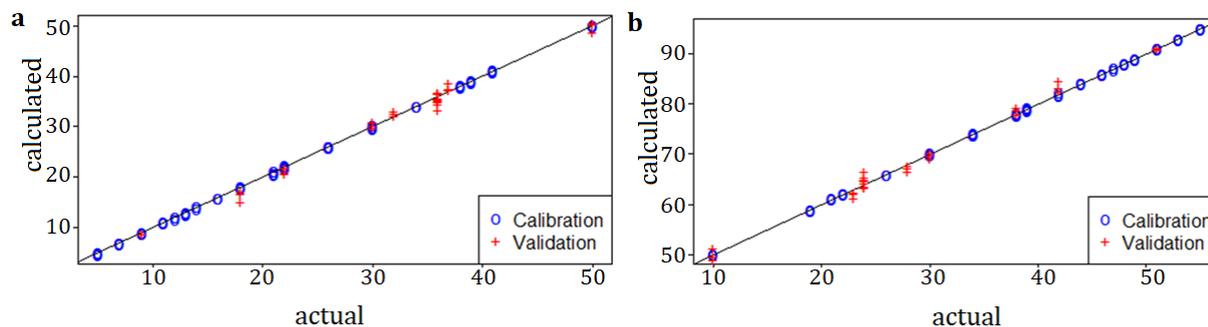


Figure 4. Prediction plots of orange oil (a) and coconut oil (b) generated from the selected calibration model for each sample

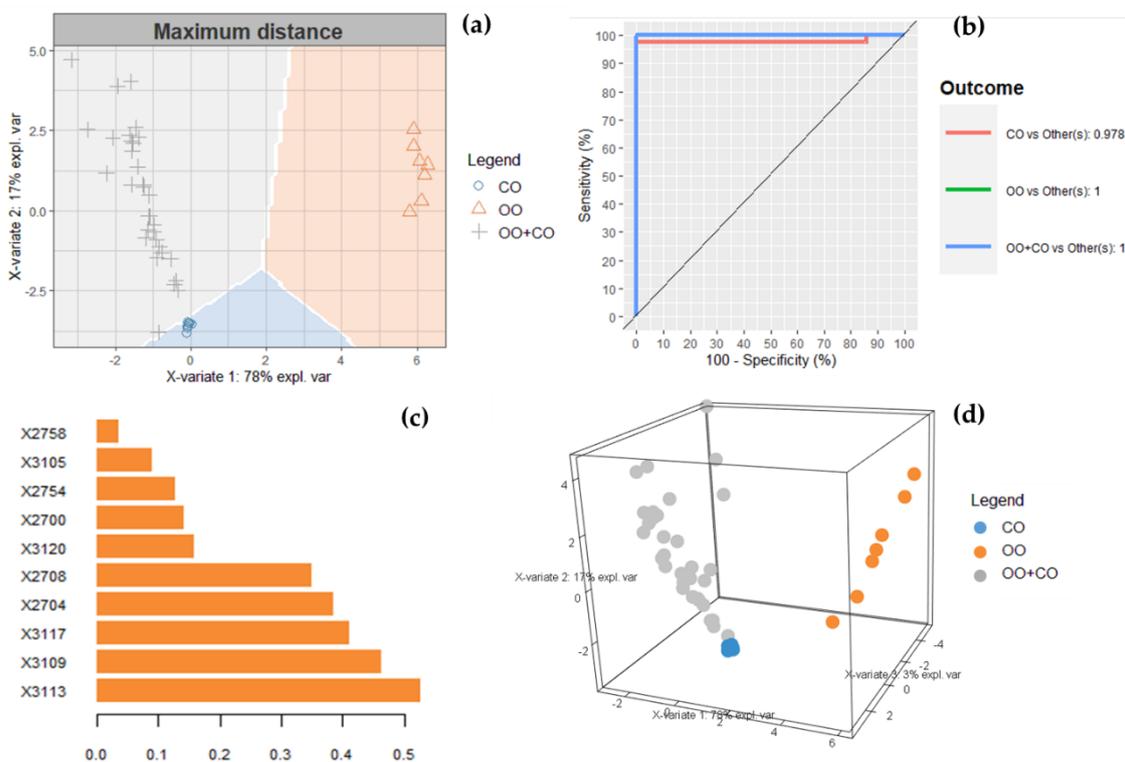


Figure 5. Sparse PLS-DA performance plots for orange oil authentication including background prediction plot (a), AUC-ROC plot (b), contribution plot (c), and 3D individual plot (d)

sPLS-DA with Tuning Parameters

The sPLS-DA, an extension to the sparse PLS, is a limited version of PLS for discrimination purposes (Lê Cao *et al.*, 2011). The main idea of applying this limited version is to combine the selection and modeling process in a one-step procedure to overcome the discrimination problem (Lê Cao *et al.*, 2008). In the case of natural product authentication, sPLS-DA

accompanied by graphical visualization is more effective than PLS-DA due to the possibility of applying the variable selection approach in multiclass problems (Jiménez-Carvelo *et al.*, 2021). This study chooses SNV spectra of OO, CO, and OO+CO to generate the sPLS-DA model since these spectra types resulted in the best multivariate calibration models for OO and CO in the previous section (Figure 5).

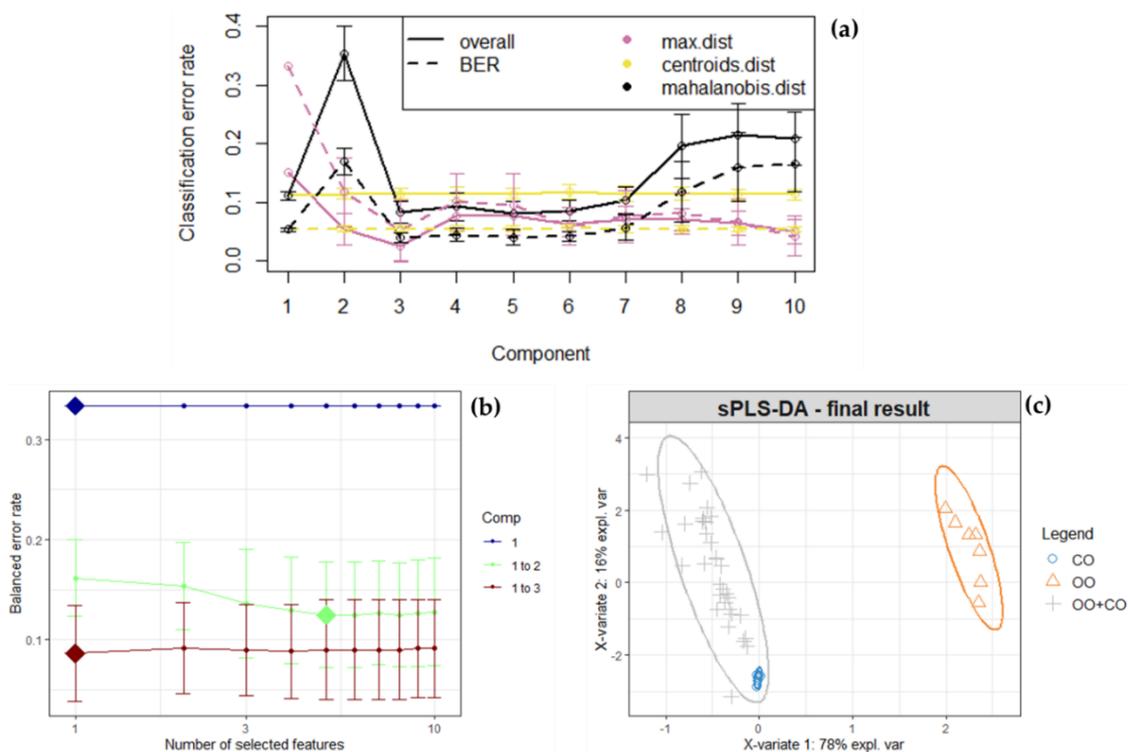


Figure 6. Sparse PLS-DA tuning parameter evaluation for classification error rates (a), features selection (b), and sPLS-DA final result (c)

The background prediction plot provides 2D visualization of the sPLS-DA model with background colors and discrimination markers for each class. The maximum distance approach is applied to classify sample categories at the predicted dummy variables with the highest dummy value (Rohart *et al.*, 2017). The AUC-ROC plot depicts the capability of the sPLS-DA model to discriminate OO, CO, and OO+CO. The AUC curve and ROC curve construct the AUC-ROC. The AUC curve represents the degree of separability of the model, while the ROC curve represents the probability of the discrimination model (Narkhede, 2018). The OO and OO+CO are separated into other classes since the model outcome shows a value of 1 or 100%. The CO outcome versus the other is 0.978. This result represented that the chance of the sPLS-DA model to distinguish CO from the others was 97.8%. The contribution plot presented the contribution of each real variable to build components for the discrimination model. The absorbance value of the SNV spectra at the wavenumber of 3113, 3109, and 3117 cm^{-1} provided the highest contribution compared to the other variables. These highest contribution

variables resulted from a narrow band above 3000 cm^{-1} . This result indicated the presence of unsaturated compounds or aromatic rings in essential oil (Agatonovic-Kustrin *et al.*, 2020; Nandiyanto *et al.*, 2019). The 3D visualization strengthens the sPLS-DA modeling to provide the model's extensive view and simplify its interpretation (Mishra *et al.*, 2021).

One of the characteristics of sPLS-DA model generation is the possibility to tune in the discrimination parameters regarding the constructed model. Aiming to apply the repetition and stratification of cross-validation for model comparison, it is crucial to determine tuning parameters and numerical outputs. For this purpose, sPLS-DA in R statistical software in the 'mixOmics' package can be implemented. The three selected parameters include component numbers, the variable of "keepX" numbers, and the prediction distance. In the cross-validation stage, this study applies 5-fold cross-validation with 30-time repetitions.

The sPLS-DA tuning parameter evaluation for classification error rates, features selection, and sPLS-DA results (Figure 6). Distance profiles of the

ten first components related to maximum distance, centroid distance, and mahalanobis distance are present along with the Balanced Error Rate (BER) and the overall value. The maximum distance approach determines the minimum classification error rate relative to the other distance. The features selection plot illustrates that the suggested list of “keepX” for components 1, 2, and 3 were 1, 5, and 1, respectively. The presentation of the sPLS-DA result follows the tuning parameters and two-dimensional plot. The X-variate 1 (x-axis) and X-variate 2 (y-axis) explained 78% and 16% of the total variances, respectively. The sPLS-DA model, with a total explained variance of 94%, successfully improves the entire explained variance compared to the previous PCA model (79.5%). In conclusion, the sPLS-DA as a supervised pattern recognition with tuning parameters proves its ability to discriminate classes of multivariate data by discarding non-informative variables during the model generation. FTIR spectral data with spectral pre-processing goes along with chemometrics techniques for developing authentication analysis for foods and herbals.

CONCLUSION

This study conducts an authentication analysis of *Citrus aurantium* L. essential oil using FTIR spectroscopy combined with chemometrics techniques. Exploratory data analysis was performed as an initial study to find the relationship between OO, CO, and OO+CO samples. Quantitative prediction models were constructed for both OO and CO in binary mixtures by applying the multivariate calibration techniques of PCR and PLSR. The study found that the PCR model using SNV spectra and the PLSR model using SNV spectra displayed the best prediction model for OO and CO, respectively. SNV spectra exploitation of OO, CO, and OO+CO built an sPLS-DA model for authentication analysis. Tuning parameters performance included component numbers, the number of “keepX” variables, and the prediction distance. With the tuning parameters approach, it successfully generated the final sPLS-DA result. Compared to the PCA model, it improved the separability of each class or category as the explained variance percentage increased. However, the applicability of the combination of FTIR spectroscopy and chemometric techniques for discriminating other types of essential oils required further investigation.

ACKNOWLEDGEMENT

This research was funded by Institute for Research and Community Services, Universitas Sanata Dharma (Research Grant Contract No. 007 Penel./LPPM-USD/II/2022). We thank Aloysia Yossy Kurniawaty for assistance with the sample collection.

REFERENCES

- Agatonovic-Kustrin, S., Ristivojevic, P., Gegechkori, V., Litvinova, T. M., & Morton, D. W. (2020). Essential oil quality and purity evaluation via ft-ir spectroscopy and pattern recognition techniques. *Applied Sciences (Switzerland)*, *10*(20), 1–12. <https://doi.org/10.3390/app10207294>
- Ali, B., Al-Wabel, N. A., Shams, S., Ahamad, A., Khan, S. A., & Anwar, F. (2015). Essential oils used in aromatherapy: A systemic review. *Asian Pacific Journal of Tropical Biomedicine*, *5*(8), 601–611. <https://doi.org/10.1016/j.apjtb.2015.05.007>
- Amit, Jamwal, R., Kumari, S., Dhaulaniya, A. S., Balan, B., & Singh, D. K. (2020). Application of ATR-FTIR spectroscopy along with regression modelling for the detection of adulteration of virgin coconut oil with paraffin oil. *LWT - Food Science and Technology*, *118*, 108754. <https://doi.org/10.1016/j.lwt.2019.108754>
- Bnina, E. Ben, Hajlaoui, H., Chaieb, I., Daami-Remadi, M., Said, M. Ben, & Jannet, H. Ben. (2019). Chemical composition, antimicrobial and insecticidal activities of the tunisian *Citrus aurantium* essential oils. *Czech Journal of Food Sciences*, *37*(2), 81–92. <https://doi.org/10.17221/202/2017-CJFS>
- Boussaada, O., & Chemli, R. (2007). Seasonal Variation of Essential Oil Composition of *Citrus Aurantium* L. var. amara. *Journal of Essential Oil-Bearing Plants*, *10*(2), 109–120. <https://doi.org/10.1080/0972060X.2007.10643528>
- Cebi, N., Taylan, O., Abusurrah, M., & Sagdic, O. (2021). Detection of orange essential oil, isopropyl myristate, and benzyl alcohol in lemon essential oil by ftir spectroscopy combined with chemometrics. *Foods*, *10*(27), 1–14. <https://doi.org/10.3390/foods10010027>

- de Araújo, J. S. F., de Souza, E. L., Oliveira, J. R., Gomes, A. C. A., Kotzebue, L. R. V., da Silva Agostini, D. L., de Oliveira, D. L. V., Mazzetto, S. E., da Silva, A. L., & Cavalcanti, M. T. (2020). Microencapsulation of sweet orange essential oil (*Citrus aurantium* var. *dulcis*) by liophylization using maltodextrin and maltodextrin/gelatin mixtures: Preparation, characterization, antimicrobial and antioxidant activities. *International Journal of Biological Macromolecules*, *143*, 991–999. <https://doi.org/10.1016/j.ijbiomac.2019.09.160>
- Derdar, H., Belbachir, M., & Harrane, A. (2019). A green synthesis of polylimonene using maghnite-H⁺, an exchanged montmorillonite clay, as eco-catalyst. *Bulletin of Chemical Reaction Engineering & Catalysis*, *14*(1), 69–78. <https://doi.org/10.9767/bcrec.14.1.2692.69-78>
- Devos, O., Downey, G., & Duponchel, L. (2014). Simultaneous data pre-processing and SVM classification model selection based on a parallel genetic algorithm applied to spectroscopic data of olive oils. *Food Chemistry*, *148*, 124–130. <https://doi.org/10.1016/j.FOODCHEM.2013.10.020>
- Djebbi, T., Soltani, A., Chargui, H., Yangui, I., Messaoud, C., & Mediouni Ben Jemâa, J. (2021). Composition and fumigant protectant potential of Tunisian *Citrus aurantium* L. essential oils against *Rhyzopertha dominica* F.(Coleoptera: Bostrichidae). *The 1st International Electronic Conference on Entomology*, 10643. <https://doi.org/10.3390/iece-10643>
- Dunning, T. (2013). Aromatherapy: Overview, safety and quality issues. *OA Alternative Medicine*, *1*(1), 1–6. <https://www.oapublishinglondon.com/article/518/>
- Dzulfiyanto, A., Riswanto, F. D. O., & Rohman, A. (2018). The employment of UV-spectroscopy combined with multivariate calibration for analysis of paracetamol, Propyphenazone and caffeine. *Indonesian Journal of Pharmacy*, *28*(4), 191–197. <https://doi.org/10.14499/indonesianjpharm28iss4pp191>
- Edris, A. E. (2007). Pharmaceutical and therapeutic Potentials of essential oils and their individual volatile constituents: a review. *Phytotherapy Research*, *21*(4), 308–323. <https://doi.org/10.1002/PTR.2072>
- Esposito, E. R., Bystrek, M. V., & Klein, J. S. (2014). An Elective Course in Aromatherapy Science. *American Journal of Pharmaceutical Education*, *78*(4). <https://doi.org/10.5688/AJPE78479>
- Gautam, R. D., Kumar, A., Kumar, R., Chauhan, R., Singh, S., Kumar, M., Kumar, D., Kumar, A., & Singh, S. (2021). Clonal Propagation of *Valeriana jatamansi* Retains the Essential Oil Profile of Mother Plants: An Approach Toward Generating Homogenous Grade of Essential Oil for Industrial Use. *Frontiers in Plant Science*, *12*(738247), 1–13. <https://doi.org/10.3389/FPLS.2021.738247/BIBTEX>
- Hemmateenejad, B., Akhond, M., & Samari, F. (2007). A comparative study between PCR and PLS in simultaneous spectrophotometric determination of diphenylamine, aniline, and phenol: Effect of wavelength selection. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, *67*(3–4), 958–965. <https://doi.org/10.1016/j.saa.2006.09.014>
- Hyldgaard, M., Mygind, T., & Meyer, R. L. (2012). Essential oils in food preservation: Mode of action, synergies, and interactions with food matrix components. *Frontiers in Microbiology*, *3*(1), 12. <https://doi.org/10.3389/FMICB.2012.00012/BIBTEX>
- Irnawati, Riswanto, F. D. O., Riyanto, S., Martono, S., & Rohman, A. (2021). The use of software packages of R factoextra and FactoMineR and their application in principal component analysis for authentication of oils. *Indonesian Journal of Chemometrics and Pharmaceutical Analysis*, *1*(1), 1–10.
- Jiménez-Carvelo, A. M., Martín-Torres, S., Ortega-Gavilán, F., & Camacho, J. (2021). PLS-DA vs sparse PLS-DA in food traceability. A case study: Authentication of avocado samples. *Talanta*, *224*(121904), 1–10. <https://doi.org/10.1016/j.talanta.2020.121904>

- Johnson, R. (2014). Food fraud and “Economically motivated adulteration” of food and food ingredients. In *Food Fraud and Adulterated Ingredients: Background, Issues, and Federal Action*.
- Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065), 1–16. <https://doi.org/10.1098/rsta.2015.0202>
- Kassambara, A. (2018). *Machine Learning Essentials: Practical Guide in R*. STHDA.
- Kassambara, A., & Mundt, F. (2017). Package “factoextra” for R: Extract and Visualize the Results of Multivariate Data Analyses. In *R package version* (pp. 1–77).
- Krishna, A., Tiwari, R., & Kumar, S. (2000). Aromatherapy-an alternative health care through essential oils. *Journal of Medicinal and Aromatic Plant Sciences*, 22(1B), 798–804. <https://eurekamag.com/research/010/201/010201750.php>
- Lê Cao, K. A., Boitard, S., & Besse, P. (2011). Sparse PLS discriminant analysis: Biologically relevant feature selection and graphical displays for multiclass problems. *BMC Bioinformatics*, 12(1), 253. <https://doi.org/10.1186/1471-2105-12-253>
- Lê Cao, K. A., Rossouw, D., Robert-Granié, C., & Besse, P. (2008). A sparse PLS for variable selection when integrating omics data. *Statistical Applications in Genetics and Molecular Biology*, 7(1), 1–25. <https://doi.org/10.2202/1544-6115.1390>
- Lee, M. Y. (2018). Essential Oils as Repellents against Arthropods. *BioMed Research International*, 2018. <https://doi.org/10.1155/2018/6860271>
- Lv, X., Zhao, S., Ning, Z., Zeng, H., Shu, Y., Tao, O., Xiao, C., Lu, C., & Liu, Y. (2015). Citrus fruits as a treasure trove of active natural metabolites that potentially provide benefits for human health. *Chemistry Central Journal*, 9(68), 1–14. <https://doi.org/10.1186/S13065-015-0145-9>
- Maksoud, S., Abdel-Massih, R. M., Rajha, H. N., Louka, N., Chemat, F., Barba, F. J., & Debs, E. (2021). *Citrus aurantium* L. Active constituents, biological effects and extraction methods. an updated review. *Molecules*, 26(19), 1–18. <https://doi.org/10.3390/molecules2619583>
- Manniche, L. (1999). *Sacred luxuries : fragrance, aromatherapy, and cosmetics in Ancient Egypt*. Cornell University Press.
- Mevik, B., & Wehrens, R. (2019). Introduction to the pls Package. In *R Package Notes*.
- Miller, J. N., Miller, J. C., & Miller, R. D. (2018). *Statistics and Chemometrics for Analytical Chemistry* (Seventh). Pearson Education Limited.
- Mishra, P., Roger, J. M., Jouan-Rimbaud-Bouveresse, D., Biancolillo, A., Marini, F., Nordon, A., & Rutledge, D. N. (2021). Recent trends in multi-block data analysis in chemometrics for multi-source data integration. *TrAC Trends in Analytical Chemistry*, 137(116206), 1–15. <https://doi.org/10.1016/J.TRAC.2021.116206>
- Moros, J., Garrigues, S., & Guardia, M. de la. (2010). Vibrational spectroscopy provides a green tool for multi-component analysis. *TrAC - Trends in Analytical Chemistry*, 29(7), 578–591. <https://doi.org/10.1016/j.trac.2009.12.012>
- Naeem, A., Abbas, T., Mohsin, T., & Hasnain, A. (2018). Brief Background and Uses. *Annals of Short Reports*, 1(1), 1–6.
- Nandiyanto, A. B. D., Oktiani, R., & Ragadhita, R. (2019). How to read and interpret ftr spectroscopy of organic material. *Indonesian Journal of Science and Technology*, 4(1), 97–118. <https://doi.org/10.17509/ijost.v4i1.15806>
- Narkhede, S. (2018). *Understanding AUC - ROC Curve*. Towards Data Science. <https://towardsdatascience.com/understanding-auc-r>
- Nurani, L. H., Rohman, A., Windarsih, A., Guntarti, A., Riswanto, F. D. O., Lukitaningsih, E., Fadzillah, N. A., & Rafi, M. (2021). Metabolite Fingerprinting Using 1H-NMR Spectroscopy and Chemometrics for Classification of Three Curcuma Species from Different Origins. *Molecules*, 26(24), 1–13. <https://doi.org/https://doi.org/10.3390/molecules26247626>
- Putri, A. R., Rohman, A., & Riyanto, S. (2019). Authentication of patin (Pangasius micronemus) fish oil adulterated with palm oil using ftr spectroscopy combined with chemometrics. *International Journal of Applied Pharmaceutics*, 11(3), 195–199. <https://doi.org/10.22159/ijap.2019v11i3.30947>

- Refaie, R., Zaghloul, S., Elbisi, M.K., & Shaaban, H. A. (2020). Bioactive Cotton for Packaging and Storage of Grains using Aromatic Components. *International Research Journal of Applied Sciences*, 2(2), 40–49. <https://scirange.com/fulltext/irjas.2020.40.49>
- Reid, L. M., O'Donnell, C. P., & Downey, G. (2006). Recent technological advances for the determination of food authenticity. *Trends in Food Science and Technology*, 17(7), 344–353. <https://doi.org/10.1016/j.tifs.2006.01.006>
- Riswanto, F. D. O., Rohman, A., Pramono, S., & Martono, S. (2021). The employment of UV-Vis spectroscopy and chemometrics techniques for analyzing the combination of genistein and curcumin. *Journal of Applied Pharmaceutical Science*, 11(3), 154–161. <https://doi.org/10.7324/JAPS.2021.110318>
- Riswanto, F. D. O., Windarsih, A., Lukitaningsih, E., Rafi, M., Fadzilah, N. A., & Rohman, A. (2022). Metabolite Fingerprinting Based on ¹H-NMR Spectroscopy and Liquid Chromatography for the Authentication of Herbal Products. *Molecules*, 27(1198), 1–17. <https://doi.org/https://doi.org/10.3390/molecules27041198>
- Rohart, F., Gautier, B., Singh, A., & Lê Cao, K. A. (2017). mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Computational Biology*, 13(11), 1–19. <https://doi.org/10.1371/journal.pcbi.1005752>
- Rohman, A., Che Man, Y. Bin, & Eakub Ali, M. D. (2019). The authentication of virgin coconut oil from grape seed oil and soybean oil using ftir spectroscopy and chemometrics. *International Journal of Applied Pharmaceutics*, 11(2), 259–263. <https://doi.org/10.22159/ijap.2019v11i2.31758>
- Rohman, A., & Che Man, Y. B. (2010). Fourier transform infrared (FTIR) spectroscopy for analysis of extra virgin olive oil adulterated with palm oil. *Food Research International*, 43, 886–892. <https://doi.org/10.1016/j.foodres.2009.12.006>
- Rohman, A., & Che Man, Y. B. (2011). Potential Use of FTIR-ATR Spectroscopic Method for Determination of Virgin Coconut Oil and Extra Virgin Olive Oil in Ternary Mixture Systems. *Food Analytical Methods*, 4(2), 155–162. <https://doi.org/10.1007/s12161-010-9156-2>
- Rohman, A., Setyaningrum, D. L., & Riyanto, S. (2014). FTIR Spectroscopy Combined with Partial Least Square for Analysis of Red Fruit Oil in Ternary Mixture System. *International Journal of Spectroscopy*, 2014, 1–5. <https://doi.org/10.1155/2014/785914>
- Sanei-Dehkordi, A., Sedaghat, M. M., Vatandoost, H., & Abai, M. R. (2016). Chemical compositions of the peel essential oil of *Citrus aurantium* and its natural larvicidal activity against the malaria vector *Anopheles stephensi* (Diptera: Culicidae) in comparison with *Citrus paradisi*. *Journal of Arthropod-Borne Diseases*, 10(4), 577–585.
- Shabaniyan, M., Hajibeygi, M., & Raeisi, A. (2020). 2-FTIR characterization of layered double hydroxides and modified layered double hydroxides. In S. Thomas & S. Daniel (Eds.), *Layered Double Hydroxide Polymer Nanocomposites* (pp. 77–101). Woodhead Publishing. <https://doi.org/https://doi.org/10.1016/B978-0-08-101903-0.00002-7>
- Sim, S. F., & Jeffrey Kimura, A. L. (2019). Partial least squares (PLS) integrated fourier transform infrared (FTIR) approach for prediction of moisture in transformer oil and lubricating oil. *Journal of Spectroscopy*, 2019(5916506), 1–10. <https://doi.org/10.1155/2019/5916506>
- Simona, J., Dani, D., Petr, S., Marcela, N., Jakub, T., & Bohuslava, T. (2021). Edible films from carrageenan/orange essential oil/trehalose—structure, optical properties, and antimicrobial activity. *Polymers*, 13(3), 1–19. <https://doi.org/10.3390/polym13030332>
- Stevens, A., & Ramirez Lopez, L. (2020). *An introduction to the prospectr package*. R Package Vignette, Report No.: R Package Version 0.1. <https://cran.r-project.org/web/packages/prospectr/vignettes/prospectr.html>
- Suntar, I., Khan, H., Patel, S., Celano, R., & Rastrelli, L. (2018). An overview on *Citrus aurantium* L.: Its functions as food ingredient and therapeutic agent. *Oxidative Medicine and Cellular Longevity*, 7864269, 1–13.
- Van Der Maaten, L. J. P., Postma, E. O., & Van Den Herik, H. J. (2009). Dimensionality Reduction: A Comparative Review. *Journal of Machine Learning Research*, 10, 1–41.

<https://doi.org/10.1080/13506280444000102>

Wińska, K., Mączka, W., Łyczko, J., Grabarczyk, M., Czubaszek, A., & Szumny, A. (2019). Essential Oils as Antimicrobial Agents—Myth or Real Alternative? *Molecules*, *24*(11). <https://doi.org/10.3390/MOLECULES24112>

130

Yang, X., Wang, C., Li, S., Huang, K., Li, M., Mao, W., Cao, S., & Xia, J. (2017). Study on the synthesis of bio-based epoxy curing agent derived from myrcene and castor oil and the properties of the cured products. *RSC Advances*, *7*(1), 238–247. <https://doi.org/10.1039/c6ra24818g>