Vibrational spectroscopy and chemometric modelling: an economical and robust quality control method for lavender oil

Tshwane University of Technology

Sidonie Tankeu, Ilze Vermaak*, Guy Kamatou, Alvaro Viljoen

Department of Pharmaceutical Sciences, Faculty of Science, Tshwane University of Technology, Private Bag X680, Pretoria 0001, South Africa * Corresponding author: vermaaki@tut.ac.za

Introduction

Lavandula angustifolia Mill. (lavender or English lavender), ranks among the top 10 medicinal plants used globally. The worldwide production of lavender oil is estimated at about 200 metric tonnes per annum which is traded extensively in the flavour and fragrance, cosmetic and aromatherapy industries [1]. René-Maurice Gattefossé was the first scientist to recognise the value of lavender oil based on his own experience of tissue regeneration on his arm which was severely burned during a laboratory explosion [2]. Since then, many studies have confirmed the use of lavender oil in the treatment of wounds, rheumatism, muscular pains, dermatitis, acne and eczema amongst many others [3]. Gas chromatography coupled to mass spectrometry (GC-MS) is the conventional method used for the quality assessment of essential oils, but this method is expensive and time-consuming. In this study, vibrational spectroscopy methods such as mid infrared (MIR) and near infrared (NIR) in combination with chemometric data analysis are proposed as alternative methods to rapidly quantify biomarkers used to determine the quality of lavender oil.

Materials and methods

GC-MS-FID analysis

Sixty commercial samples of lavender oil were analysed using gas chromatography coupled to a mass spectrometer and flame ionisation detector (GC–MS–FID). The Agilent 6860 N GC system was coupled directly to a 5973 MS to obtain reference quantification data. All the samples (diluted to a concentration of 20% in hexane) were injected (1 µl) with a (200:1) split ratio at 24.79 psi and an inlet temperature of 250°C. Carrier gas: helium; flow rate: 1.2 ml/min; electron impact 70 eV; and a scanning range of 35 to 450 *m/z*. Percentages were obtained by electronic integration measurement of peak areas, using FID. Compound identification was carried out using standard libraries (NIST and Mass Finder).

MIR spectroscopy

A volume of 10 µl of each essential oil sample was use to cover the diamond crystal surface of an Alpha-*P* Bruker spectrometer. OPUS® software was used to acquire duplicate MIR spectra of 32 scans in absorbance mode within a wave range of 550–4000 cm^{-1}. The average was calculated using Microsoft® Excel and chemometric data analysis was performed using SIMCA-P^+ 12.0 software (Umetrics AB, Malmo, Sweden).

NIR spectroscopy

The NIR Flex N500 spectrometer equipped with a liquid cell was used to acquire NIR spectra. High precision cuvettes of 0.20 mm path length were filled with 50 µl of sample. Duplicates of 32 scans per sample were collected in transmittance mode with the spectral resolution set at 4 cm^{-1} in the range of 10 000–4000 cm^{-1}. Data were converted to absorbance and chemometric analysis performed on the averages using SIMCA-P^+ 12.0 software (Umetrics AB, Malmo, Sweden).

Data analysis

Principal component analysis (PCA) was performed with different scaling methods including univariate, center and pareto. Score scatter plots were used to identify strong outliers and the correlation among the Y-variables was determined using PCA-Y. Regression of the spectral data against the GC reference values was performed to construct calibration models for six compounds including 1,8-cineole, (*E*)-β-ocimene, (*Z*)-β-ocimene, camphor, linalool, and linalyl acetate. Each constructed model represented a specific compound for both MIR and NIR data. Two samples were randomly selected using SIMCA-P^+ 12.0 software for external model validation and the remaining 58 samples were used for calibration. The models were tested in five random rounds for all the major compounds and the models with the best statistical performance were considered in terms of the coefficient of determination (R^2), the Q^2 value or coefficient of prediction, root mean square error of prediction (RMSEP) and the root mean square error of estimation (RMSEE) [4]. Finally, the external dataset consisting of two samples were predicted using the best calibration model constructed for each individual compound. The reliability and accuracy of the model was assessed by comparing the predicted values to the reference values obtained from GC–MS–FID analysis of the oils.

Table 1. Prediction values of the two unknown samples using MIR and NIR PLS calibration models

Compounds	S	MIR predictions (%)	GC-MS data (%)	NIR predictions (%)
1,8-Cineole	S1	4.6	4.7	4.9
	S2	7.2	7.2	6.7
(*E*)-β-Ocimene	S1	1.0	0.5	1.1
	S2	1.3	1.8	1.2
(Z)-β-Ocimene	S1	0.8	0.2	0.8
	S2	1.8	3.2	0.9
Camphor	S1	5.6	6.5	5.7
	S2	8.6	8.5	7.2
Linalool	S1	37.3	37.2	36.1
	S2	34.1	34.7	33.6
Linalyl acetate	S1	30.4	30.0	30.9
	S2	28.2	26.0	28.9

S-sample; MIR-mid infrared; NIR-near infrared; GC–MS-gas chromatography coupled to mass spectrometry.

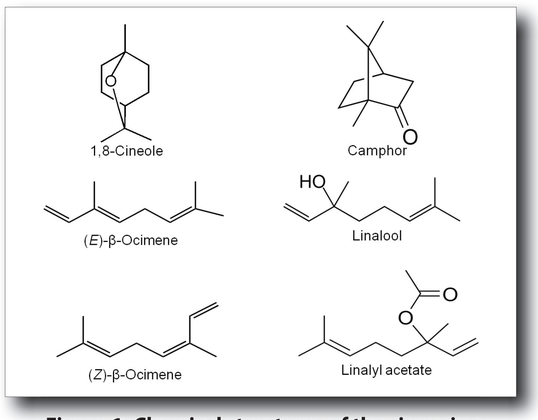

Figure 1: Chemical structures of the six major compounds quantified in lavender oil.

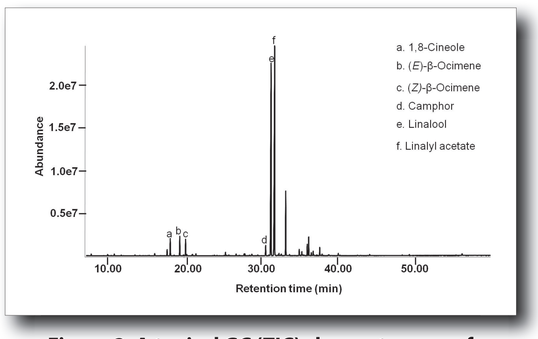

Figure 2: A typical GC (TIC) chromatogram of lavender oil indicating the six compounds included in the chemometric modelling.

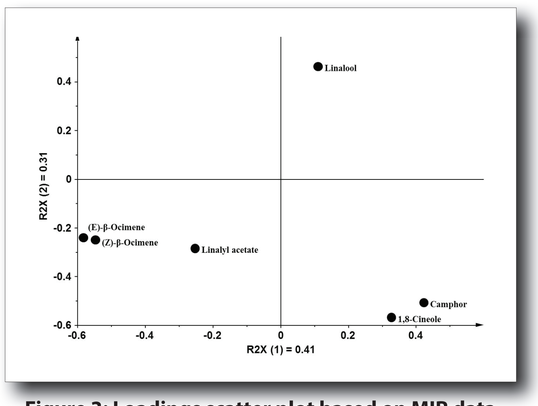

Figure 3: Loadings scatter plot based on MIR data showing the correlation among the six compounds.

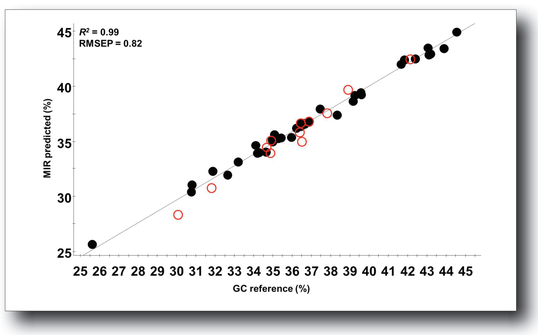

Figure 4: PLS plot of predicted MIR values versus GC–MS reference values for linalool. 70% of calibration set (●) and 30% of validation set of the samples (○) was randomly selected.

Results and discussion

GC–MS–FID analysis

The chemical structures of camphor, 1,8-cineole, linalool, linalyl acetate, (*E*)-β-ocimene and (*Z*)-β-ocimene quantified from lavender oil are shown in Figure 1. An example of a typical GC total ion chromatogram (TIC) for lavender oil is shown in Figure 2. Linalool and linalyl acetate represented the highest percentage composition of the oil and were present in quantities with a range of 25.7–45.5% for linalool and 10.2–45.6% for linalyl acetate. Large variation was observed in the chemical composition and this is advantageous as the calibration model developed can be used to predict the composition of a wide variety of samples.

Principal component analysis

First derivative pre-treatment with centre scaling produced the best results and was selected to construct the calibration models. The correlation among the six major compounds was visualised by plotting PCA-Y as shown in Figure 3. (*E*)-β-Ocimene, (*Z*)-β-ocimene and linalyl acetate were negatively correlated; inversely to linalool. A direct reaction can produce linalyl acetate from linalool through an acetyl-coenzyme A reaction while (*E*)-β-ocimene can be reduced to linalool; these biosynthetic sequences could explain the correlation among the compounds. The two major compounds, linalool and linalyl acetate which are inversely correlated, are responsible for the biological activity and therapeutic effect of lavender oil.

Partial least squares calibration models

Good coefficients of determination (R^2) of ≥0.82 were obtained with MIR and NIR data for the six major compounds. The RMSEP and RMSEE values where generally low (≤1.6) for both MIR and NIR data with the RMSEE lower than the RMSEP. The (*E*)-β-ocimene, linalool, and linalyl acetate calibration models showed very good R^2 values of ≥0.90 based on MIR as well as NIR data. The best coefficients of determination were for linalool with an R^2 of 0.98 for NIR and 0.99 for MIR data. An example of a calibration curve based on MIR spectral data for linalool is shown in Figure 4.

Prediction of the external dataset

The calibration models of the six major compounds were used to predict the composition of two external lavender oil samples and these values were compared to the reference GC–MS–FID data to determine the accuracy. Table 1 presents the prediction data as well as the reference data. The statistical accuracy of the predictions was evaluated using model membership probability values (PModXPS+). Almost all of the values were greater than 5% as recommended by Eriksson *et al*. [4]. Other important parameters are the distance to the model (DModXPS+) and the Dcrit which is the critical distance line/value indicating whether predictions are within the model's critical limit. The values of these parameters demonstrated the accuracy of predictions of the external samples [4].

Conclusions

The results obtained indicated that calibration models based on both MIR and NIR spectral data could be successfully used to predict major and several minor compounds present in unknown lavender oil samples based on the statistical parameters. The prediction accuracy of major lavender oil compounds was higher compared to the predictions for the minor compounds and the MIR spectral data provided superior results in comparison to the NIR spectral data. The results demonstrated that MIR and NIR spectroscopy can be implemented as rapid, less expensive, efficient, chemical free, and non-destructive alternative methods in the quality control of lavender oil for routine analysis applicable to industry.

Additional information at: http://dx.doi.org/10.1016/j.indcrop.2014.05.005

Acknowledgements

Tshwane University of Technology and the National Research Foundation (NRF) for the financial support.

References

[1] DAFF (Department of Agriculture, Forestry and Fisheries), 2009. *Lavender Production*. Directorate Agricultural Information Services, South Africa, pp. 1–20.

[2] Vincent, E., 2008. Lavender Essential Oil, Available at: http://www.youngliving.com/enUS/products/essential-oils/singles/lavender-essential-oil [accessed on: 2013/01/20].

[3] Van Wyk, B-E., Wink, M., 2004. *Medicinal Plants of the World*, first edition. Briza Publications, Pretoria, p. 480.

[4] Eriksson, L., Johansson, E., Kettaneh-Wold, N., Trygg, J.C., Wikstrom, C., Wold, S., 2006. *Multi- and Megavariate Data Analysis Part 1: Basic Principles and Applications*, second edition. Umetrics AB, Sweden, pp. 385–386.