Journal of Food Engineering 364 (2024) 111795

Contents lists available at ScienceDirect

Journal of Food Engineering

journal homepage: www.elsevier.com/locate/jfoodeng

Rapid determination of the storage time of cold-pressed berry seed oils using flash gas chromatography E-Nose coupled with chemometrics

Yolanda Victoria Rajagukguk [a], Chiara Cevoli [b,*], Ilaria Grigoletto [b], Jolanta Tomaszewska-Gras [a]

[a] Department of Food Quality and Safety Management, Poznań University of Life Sciences, ul. Wojska Polskiego 31/33, 60-624, Poznań, Poland
[b] Department of Agricultural and Food Sciences (DISTAL), University of Bologna, Cesena, Piazza Goidanich 60, 47521, Italy

ARTICLE INFO

Keywords:
Oil deterioration
Authenticity
Non-targeted methods
Chemometrics
Storage
FGC E-nose
Volatile compounds

ABSTRACT

Since oils containing a high content of polyunsaturated fatty acids are susceptible to oxidation, it is necessary to monitor the degree of deterioration during storage, e.g. by measuring the volatile compounds. This study aimed to assess volatile profiles of berry seed oils in terms of the authenticity and the deterioration assessment using flash gas chromatography (FGC E-Nose) combined with chemometrics. Berry seed oils (raspberry, blackcurrant, strawberry, chokeberry), obtained from three different suppliers and stored for a one year in brown bottles at room temperature, were analysed after 0, 3, 6, 9 and 12 months of storage. Principal component analysis enabled separation of oil samples by different berry types, suppliers and storage times. To predict the storage time, partial least square (PLS) models were built for each type of berry oil. Determination coefficients (R^2) in cross-validation ranged from 0.842 (RMSECV = 1.69 months) to 0.969 (RMSECV = 0.75 months). Selecting the specific regions of chromatograms improved the residual prediction deviation (RPD) to values between 2.8 and 5.7, which indicated the suitability of the PLS models to predict the storage time in the quality control of berry oils.

1. Introduction

World production of berries was increased during the past 10 years. The growth in production was observed in the increment of gross production value that ranged up to 25% and 125%, respectively, for strawberry and raspberry commodities (FAO, 2021). Berries were often consumed directly or processed into other convenient products (juice, dried berries, confiture, and marmalade). Agro-industrial processing resulted to the accumulation of by-products such as pomace and other residues (seeds, leaves, stems) that account up to 35% of the raw mass (Majerska et al., 2019). Furthermore, up to 70 % of berry by-products consist of seeds containing a notable amount of oil (Mazurek et al., 2022). Oils from berry seeds are characterized by a high percentage of polyunsaturated fatty acid (PUFA) and antioxidant components, which are linked to health-promoting properties, such as reducing the risk of cardiovascular disease (Martysiak-Żurowska and Orzołek, 2023).

Berry seed oils are sold as food supplements, nutraceutical products, and cosmetics. Despite the growing demand, only a limited number of studies exist on berry seed oils' authenticity (Przykaza et al., 2021; Rajagukguk et al., 2023). The quality of berry seed oil sold on the market remains unprotected due to the non-existing regulation of oils from by-products. The high profitability of berry seed oil production might attract fraudulent enterprises to tamper with the oils. In fact, fat and oil products remain the third-most reported commodities in terms of the suspicion of fraud in the latest report from the EU-Food Fraud Network, with "faulty storage conditions" and "unsuitable organoleptic characteristics" included in the list of non-compliance categories (European Commission, 2021). From the industrial point of view, developing a reliable method for evaluating storage time is critical to ensure oil quality and to protect the business-to business marketplace against fraud cases.

During storage, oils with higher PUFA are more susceptible to oxidation. The deterioration in quality is mainly indicated by the presence of rancid odours from lipid oxidation products such as aldehydes, ketones, and esters, as well as furan derivatives (Gaca et al., 2021). Quality and authenticity assessment of oils was reported in recent studies using various methods of gas chromatography techniques (Mota et al., 2021). Due to its simplicity, efficiency, and affordability, the flash gas chromatography (FGC) E-nose application has gained attention as a non-targeted instrument to assess the quality of high-lipid-containing products (Barbieri et al., 2020; Cevoli et al., 2022; Tata et al., 2022). FGC E-Nose is a versatile tool that enables rapid analysis of the samples

* Corresponding author.
E-mail address: chiara.cevoli3@unibo.it (C. Cevoli).

https://doi.org/10.1016/j.jfoodeng.2023.111795
Received 17 July 2023; Received in revised form 25 September 2023; Accepted 13 October 2023
Available online 14 October 2023

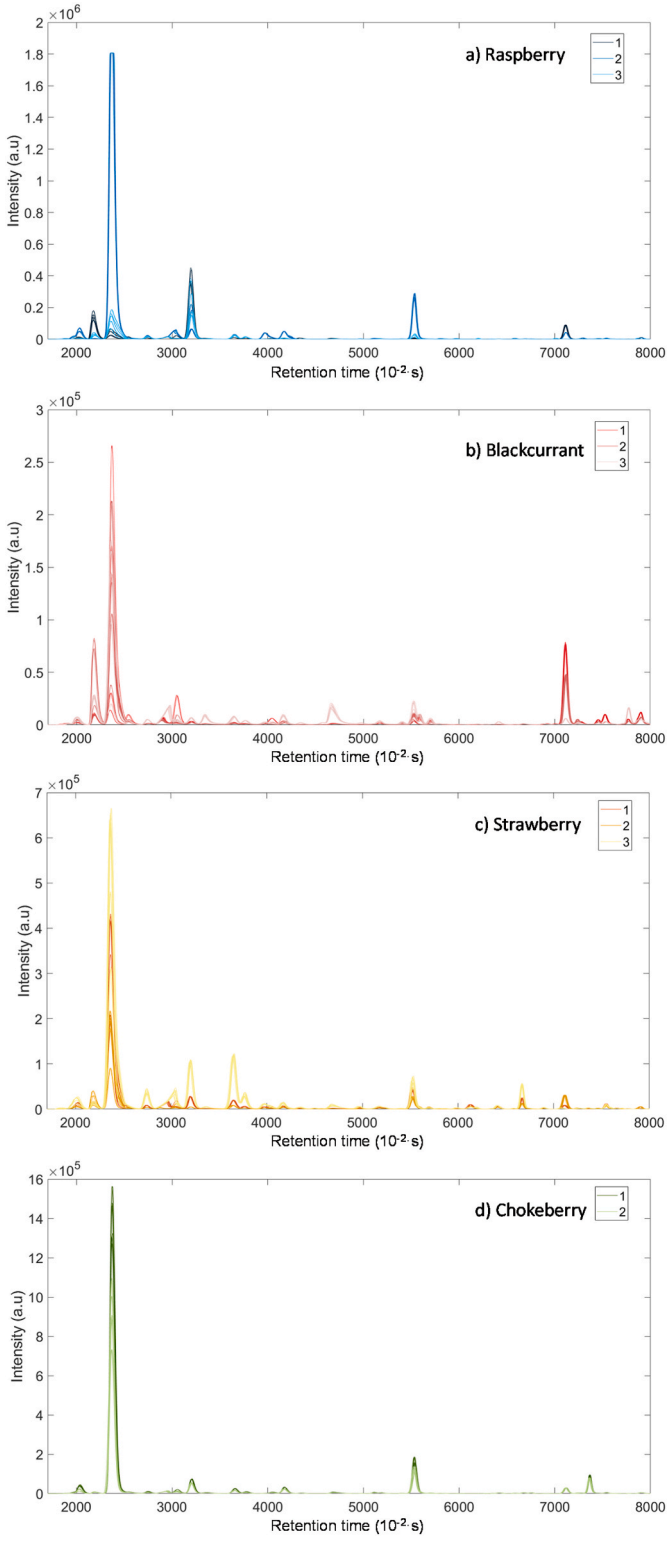

Fig. 1. Chromatogram of a) raspberry, b) blackcurrant, c) strawberry, d) chokeberry seed oils from different suppliers.

compared to conventional GC. The synergy between FGC and E-Nose instruments leads to better data acquisition, i.e. enabling researchers to obtain more specific and complex volatile signals from flavour attributes compared to sensor-based E-nose (Tian et al., 2023). Furthermore, FGC is easy to operate and suitable for profiling volatile compounds, which can be evaluated with chemometric analysis (Damiani et al., 2020). The implementation of FGC E-Nose is still a hot topic in edible oil studies. The technique capabilities were proven in determining olive oil authenticity (Palagano et al., 2021), the origin of flavoured rapeseed oil (Zhang et al., 2020), and has even been encouraged for use as a replacement for conventional sensory evaluations (Modesti et al., 2021). Further investigation of the implementation of FGC E-Nose to assess edible oil quality will benefit the oil industry.

The objectives of this study were: a) to predict the storage time of berry seed oils using FGC E-nose combined with multivariate data analysis and, b) to discriminate four types of berry seed oils obtained from different suppliers based on the oils' authentic volatile profiles. The samples used in this study consist of cold-pressed oil from raspberry (RB), strawberry (ST), blackcurrant (BC), and chokeberry (CHB). Chromatogram data will be subjected to multivariate data analysis to predict the storage time (PLS model) and to assemble the samples according to berry type and supplier origin (PCA). This work presented the first application of FGC E-Nose to evaluate berry seed oil quality in a storage study.

2. Material and methods

2.1. Materials

Four types of cold-pressed berry seed oils were analysed: raspberry (*Rubus idaeus*) [RB], strawberry (*Fragaria ananassa*) [ST], blackcurrant (*Ribes nigrum*) [BC], and chokeberry (*Aronia melanocarpa*) [CHB]. For each oil type, three different batches of production (GR1, GR2, and GR3) coming from two supplier (GreenField Sp z.o.o., Warsaw, Poland; Olvita Gołuch Sp. k., Marcinowice, Poland), were investigate, except that for the chokeberry seed oil. In this last case two batches were considered. Freshly produced oils were stored for a one-year observation in brown-glass bottles (100 ml) without any headspace at ±20 °C. The samples were collected from different bottles every 3 months. In total, 110 samples were collected (two replicas for storage time) and analysed after 0, 3, 6, 9, and 12 months of storage.

2.2. Flash gas chromatography (FGC) analysis

Non-targeted volatile compounds analysis was conducted using a flash gas chromatography technique and the Heracles II E-Nose instrument (Alpha MOS, Toulouse, France). FGC is equipped with two capillary columns (10 m length, 180 μm diameter) containing different stationary phase polarities. Non-polar (MXT5: 5% diphenyl, 95% methylpolysiloxane) and polar (MXT-1701 14% cyanopropylphenyl/86% dimethyl polysiloxane) columns were parallelly separated into compounds with a distinctive capability. The separated volatile compounds were detected by flame ionisation detector (FID).

Berry seed oils (±2 g) were weighed inside a 20 mL clear-glass vial and sealed with a magnetic cap. Each sample was analysed in two replications. The vials were placed inside a shaker oven for 20 min (40 °C, 500 rpm). Volatile compounds (5 mL) were collected from the vial's headspace using syringe (70 °C) equipped with a splitless injector (200 °C, injection speed 100 μL/s, carrier gas flow at 30 mL/min) and a Tenax TA® trap (40 °C, 60 s) as the adsorbent. Afterwards, volatile compounds were desorbed from the trap (240 °C, 93 s), injected (column head pressure was set at 40 kPa), and split into two columns (flow 5 mL/min). The analysis was conducted using the following thermal program: 1) start at 40 °C and hold for 2 s, 2) increase the temperature to 80 °C at 1 K/s, 3) increase the temperature to 250 °C at 3 K/s. The carrier gas used was hydrogen (40 kPa–64 kPa, with increasing rate of 0.2 kPa/s). The compounds were detected by a flame ionisation detector (FID) at 260 °C. Alphasoft software version 14.5 was used to record the signals and control the whole process of data acquisition (signal digitalized every 0.01 s).

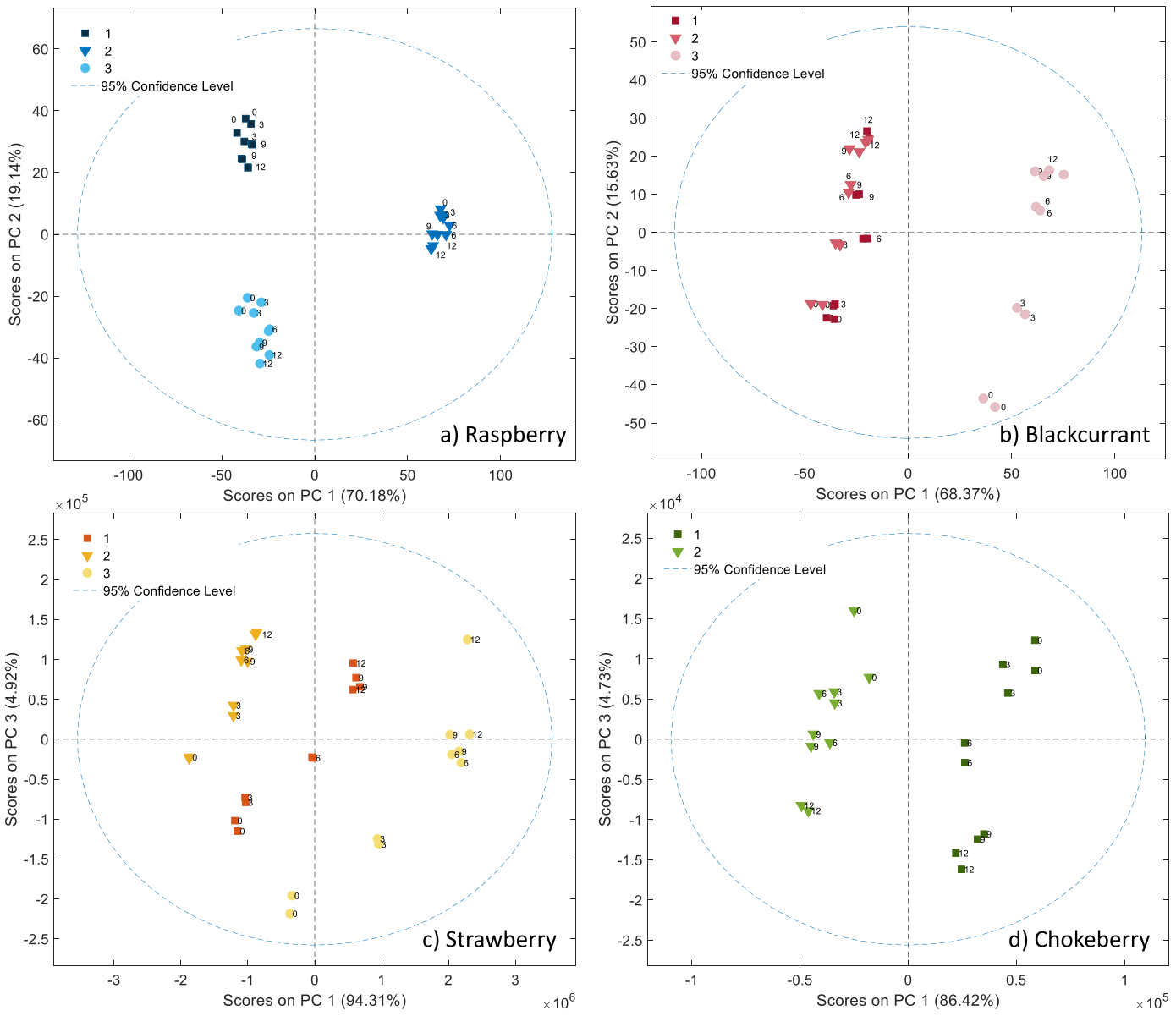

Fig. 2. PCA score plots for a) raspberry, b) blackcurrant, c) strawberry, d) chokeberry seed oils during 0, 3, 6, 9, 12 months of storage.

2.3. Data analysis

By using a non-targeted approach, full chromatograms or some sections of them were used to predict the storage time and to discriminate the sample according to berry types and batch origin. The choice of using full chromatograms has some advantages related to possible errors generated during the integration in peak-area calculation; furthermore, when no pre-selection is done, the risk of discarding useful information is avoided (Melucci et al., 2016). Before the chemometric analysis, the chromatograms belonging to the same berry types were aligned by the COW (Correlation Optimized Warping) algorithm (Tomasi et al., 2004) and pretreated using two methods: Pareto-scaling or simply centering (mean-centering). Pareto-scaling was selected instead of standard auto-scaling because it gives equal importance to all variables, but to a lesser extent than standard autoscaling, which may induce a loss of significant information, since it increases the weights of minor noisy variables (Aliakbarzadeh et al., 2016). Furthermore, by using this pre-processing, the shapes of the chromatograms were not modified (van den Berg et al., 2006). Particularly, in Pareto-scaling the square root of the standard deviation is used as scaling factor:

$$\widetilde{x}_{ij} = \frac{x_{ij} - \overline{x_i}}{\sqrt{SD_i}} \tag{1}$$

where x_{ij} and $\widetilde{x}_{ij}$ represent the raw and scaled data, respectively. The mean ($\overline{x_i}$) and standard deviation (SD_i) are calculated as:

$$\overline{x_i} = \frac{1}{J}\sum_{j=1}^{J} x_{ij} \tag{2}$$

$$SD_i = \sqrt{\frac{\sum_{j=1}^{J}\left(x_{ij} - \overline{x_i}\right)^2}{J-1}} \tag{3}$$

Explorative PCA models were built to visualize samples according to berry type and batch origin, while PLS models were developed to estimate the storage time within the same berry type, considering all suppliers together. In view of the small number of samples, the PLS model validation were performed by Venetian blinds cross-validation (10 segments). To avoid the model over-fitting, the optimal number of latent variables (LV) were chosen by detecting the global minimum of root

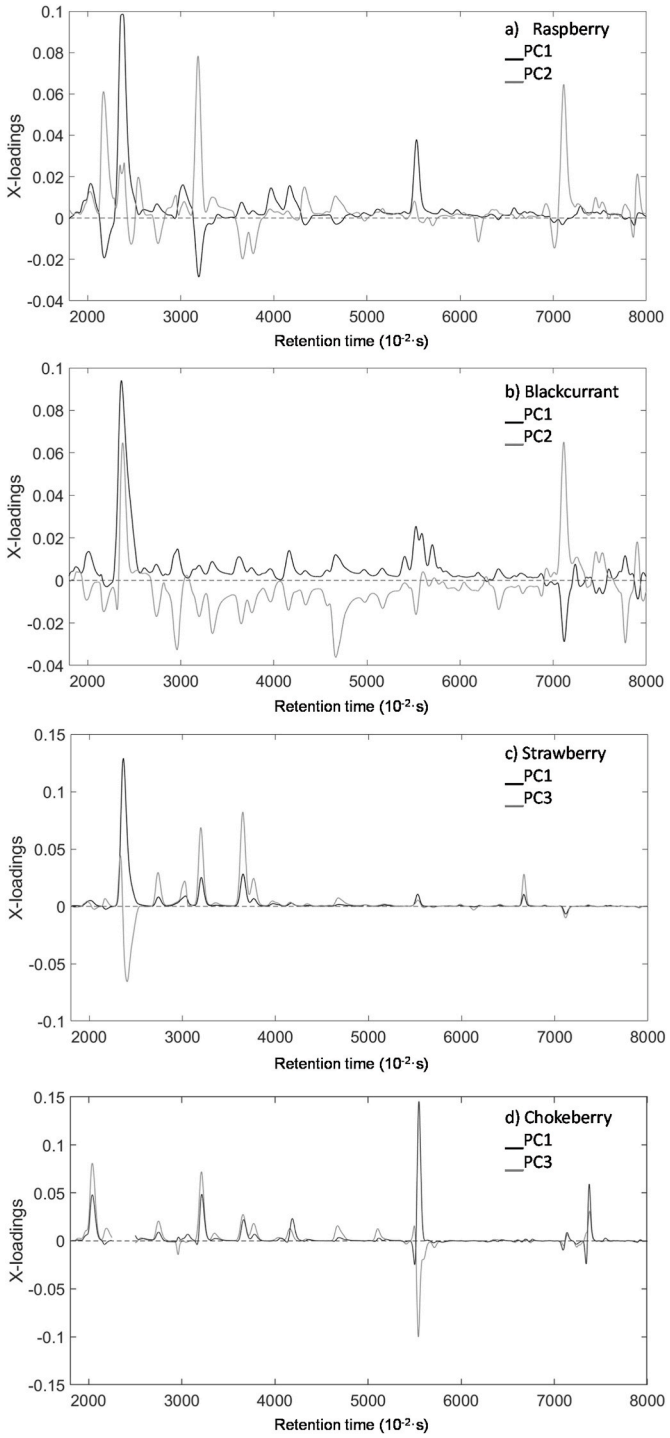

Fig. 3. X-loadings score plots obtained by the PCA of the a) raspberry, b) blackcurrant, c) strawberry, d) chokeberry data.

mean square error in cross validation (RMSECV).

Furthermore, the model robustness and significance was verified through a permutation test. This procedure allows overfitting to be identified and provides the probability that the given model is significantly different from one built under the same conditions but using random data.

It is often the case that most of the variables involved in the PLS model development could be of slight relevance (e.g. redundant or unnecessary chromatogram regions) to the investigated problem, as they represent variation not related to the response to be modelled.

Table 1
Results of the PLS models developed using a full chromatogram (1800–8000).

Berry	Pre-processing	Calibration		Cross-validation			LV
		R^2	RMSEC (months)	R^2	RMSECV (months)	RPD	
RB	PA + MC	0.969	0.75	0.842	1.69	2.5	8
BC	PA + MC	0.984	0.54	0.846	1.63	2.5	11
ST	MC	0.985	0.5	0.932	1.14	3.8	11
CHB	PA + MC	0.995	0.26	0.963	0.82	5.2	8

RB: raspberry; BC: blackcurrant; ST: strawberry; CHB: chokeberry; PA: Pareto Scaling, MC: mean-centering; R^2: determination coefficient; RMSE: Root Mean Square Error; C: calibration; CV: cross-validation; RPD: residual prediction deviation; LV: latent variables.

Therefore, their number can be drastically reduced without loss of information, or even increasing the model power. Variable selection can improve the estimation accuracy by effectively identifying the subset of important predictors and can enhance the model's interpretability with parsimonious representation (Farrés et al., 2015). There are several methods to select variables, and one of those most used in combination with PLS regression is the Variable importance in projection (VIP) selection method. VIP scores summarize the influence of individual X-variables on the PLS model. They are calculated as the weighted sum of squares of the PLS weights, which take into account the amount of explained Y-variance in each extracted latent variable (component).

The VIP score for each j variable is calculated as:

$$\mathrm{VIP_j} = \sqrt{\frac{\sum_{f=1}^{F} w_{jf}^2 \bullet \mathrm{SSY_f} \bullet \mathrm{J}}{\mathrm{SSY_{tot}} \bullet \mathrm{F}}} \tag{4}$$

where w_{jf} is the weight value for j variable and f component, $\mathrm{SSY_f}$ is the sum of squares of explained variance for the fth component and J number of X variables. $\mathrm{SSY_{tot}}$ is the total sum of squares of explained variance of the dependent variable, and F is the total number of components.

X variables characterized by VIP scores higher than 1 are considered important in a given model; this criterion is conventionally used to select the variable. Accordingly, new PLS models involving only the X variables with VIP scores greater than one were developed.

The results were evaluated in terms of determination coefficient (R^2), root mean square error in (RMSE), and residual prediction deviation (RPD) in calibration (C) and validation (CV)

$$R^2 = \frac{\sum_{i=1}^{N} (\widehat{y_i} - \overline{y})^2}{\sum_{i=1}^{N} (y_i - \overline{y})^2} \tag{5}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (\widehat{y_i} - \overline{y})^2}{N}} \tag{6}$$

$$RPD = \frac{SD}{RMSE} \tag{7}$$

where y_i is the actual storage time (days), $\widehat{y_i}$ is the predicted storage time (days), $\overline{y}$ is the mean of the actual values, N is the number of samples, and SD is the standard deviation of reference values. All data analyses were conducted using PLS Toolbox for Matlab2018a®.

3. Results and discussion

All chromatograms, grouped based on the berry type, are shown in Fig. 1. The main peaks are concentrated in the initial part of the chromatogram (retention time between 18 and 80 s). As could be expected, the volatile profiles of the four types of oils (RB, BC, ST and CHB) are quite different, in terms of shape and peak numbers. Clear differences between supplier origin can be observed for all berry types, especially in

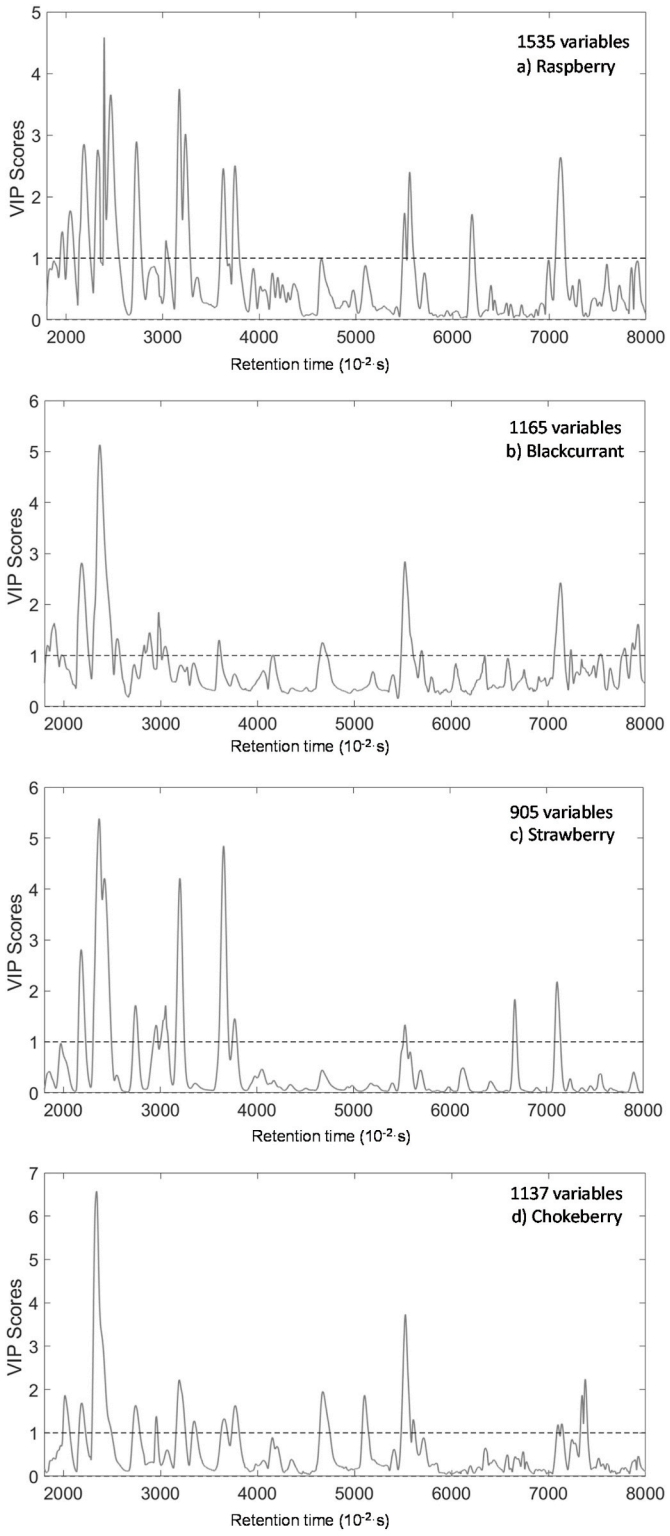

Fig. 4. VIP scores from the selected regions in chromatogram for a) raspberry, b) blackcurrant, c) strawberry, d) chokeberry seed oils.

terms of intensities, confirming the discriminating power of the volatile profile with respect to the supplier. This draws attention to the fact that sample oils obtained from two different batches of production from the same supplier (group 1 and 2) are characterized by different volatile profiles, even immediately after production.

Explorative PCAs were conducted to evaluate a possible correlation between chromatograms according to berry type and supplier origin. Score plots obtained for the chromatogram range from 1800 to 8000 points (retention time from 18 to 80 s) are shown in Fig. 2. Better results were achieved for pre-processing RB and BC data by Pareto-scaling, and for ST and CHB data by mean-centering. For all berry types, a good separation was observed for the supplier origin PC1 *vs* PC2 (RB and BC) or PC1 *vs* PC3 (ST and CHB). For the BC samples, it is difficult to discriminate between group 1 and 2, while group 3 shows high differences with respect to the other two. This is probably due to the fact that samples belonging to group 3 are produced by a different company.

Within the same supplier, it is also possible to observe quite good distribution of the samples as a function of storage time, from 0 to 12 months, especially along the PC2 (RB and CB) or PC3 (ST and CHB).

The contribution of X-variable to each of the PCs can be evaluated by the X-loading. High loading values (positive or negative) indicate that a variable has a strong effect on that principal component. Positive loadings indicate a positive correlation between variable and principal component: an increase in one results in an increase in the other. Negative loadings indicate a negative correlation. Therefore, by evaluating the X-loadings (Fig. 3), it was possible to observe the chromatographic zones characterized by the highest contribution to PC1 *vs* PC2 (RB and CB) or PC1 *vs* PC3 (ST and CHB). The presence of greater noise in the loading scores of the RB and CB samples compared to the ST and CHV, is due to the data pre-treatment (Pareto-scaling *vs* mean centring).

A brief attempt to identify the FGC E-Nose peaks from olive oils of different geographical origin was reported in the study conducted by Melucci et al. (2016). The authors compared the FGC E-Nose peaks with MS spectra in SPME/GC-MS analysis and successfully identified a positive correlation between retention times and particular molecules. According to Melucci et al. (2016), a rough identification of the responsible molecules for some prominent peaks from FGC E-Nose in Fig. 1 can be established. The identified peaks were listed as follows: a) ethyl acetate at 2100, b) ethanol 2200, c) 1-penten-3-ol at 2400, d) hexanal at 3200, e) hexanol at 3600, and f) 2-hexenal at 4100.

Considering the discrimination according to the storage time (observed along PC2 or PC3), the highest contribute is mainly due to the peaks at around 2200, 3200 and 7100; 2400 and 7100; 2400, 3200 and 3600; as well as 2100, 2400, 3200 and 5600, for RB, BC, ST and CHB, respectively. In this regards, several prominent volatile compounds that responsible for storage time differentiation as detected by FGC E-Nose are 1-penten-3-ol, hexanal, and 2-hexenal. Hexanal and 2-hexenal were generated from the oxidation of linoleic and linolenic acid respectively (Xu et al., 2018), which responsible for up to 75% of the fatty acid composition in berry seed oils (Mildner-Szkudlarz et al., 2019). While the increment of 1-penten-3-ol over time indicates the decomposition of hydroperoxides from omega-3 fatty acid in berry seed oils (Liang et al., 2020). Beside quality deterioration and off-flavours, the detected volatiles were also responsible for the key-aromas that set apart one type of berry seed oil among the other oils. For example 1-hexanol (fruit, banana, soft, tomato, cut grass), 2-hexenal (green, apple-like, bitter almond like), and hexanal (apple, cut grass, green) (Marx et al., 2021; Teixeira et al., 2021). Regardless the detected peaks in the presented chromatogram, it should be noted that FGC E-Nose is a completely non-targeted method. Hence, a non-targeted approaches was used to elaborate the data, consequently the precise identification of the volatile compounds related to the peaks is outside the scope of the work. However, this study does not exclude the fact that further studies could be done on the analytical determination of these volatile compounds from a purely chemical point of view.

PLS models were developed to estimate the storage time (months) within the same berry type, regardless of the supplier origin, by using full chromatograms (from 1800 to 8000 ponits). Model results, in terms of determination coefficient (R^2), RMSE, RPD and (LV) in calibration and cross-validation (10 segments) are shown in Table 1. Good results were obtained for all berry oils with R^2 in cross-validation ranging from 0.842 (RMSECV = 1.69 months, RB) to 0.963 (RMSECV = 0.82 months, CHB), while RPD values are equal to or higher than 2.5 (up to 5.2 for

Table 2
Results of the new PLS model developed using a selected region in the chromatogram by the VIP method.

Berry	Pre-processing	Numbers of variable	Calibration		Cross-validation			LV
			R^2	RMSEC (months)	R^2	RMSECV (months)	RPD	
RB	PA + MC	1535	0.980	0.60	0.973	1.01	4.0	5
BC	PA + MC	1165	0.970	0.73	0.932	1.14	3.8	7
ST	MC	905	0.973	0.68	0.895	1.42	3.1	4
CHB	PA + MC	1137	0.989	0.44	0.969	0.75	5.7	5

RB: raspberry; BC: blackcurrant; ST: strawberry; CHB: chokeberry; PA: Pareto Scaling, MC: mean-centering; R^2: determination coefficient; RMSE: Root Mean Square Error; C: calibration; CV: cross-validation; RPD: residual prediction deviation; LV: latent variables.

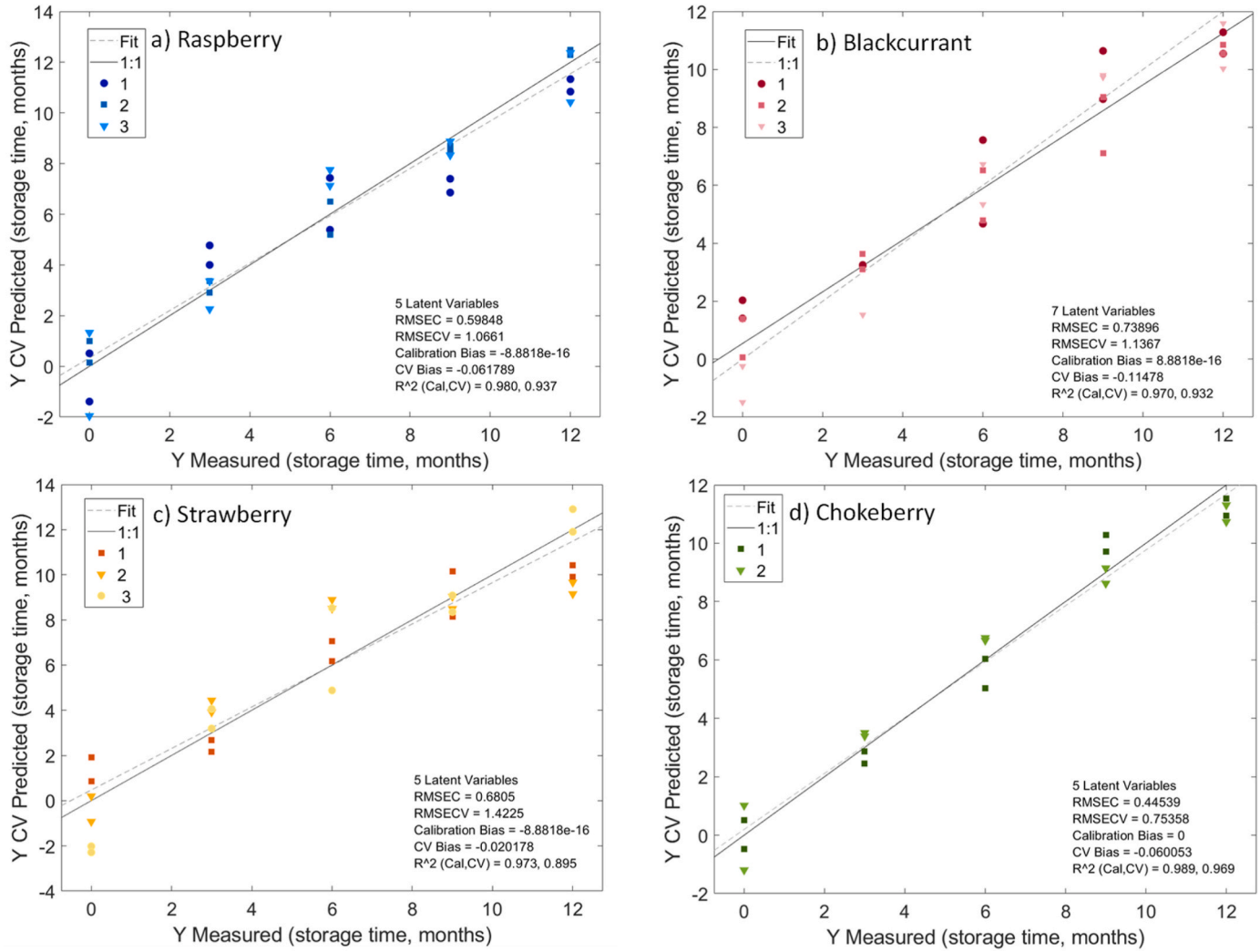

Fig. 5. PLS results in terms of measured versus predicted storage (time months) values for a) raspberry, b) blackcurrant, c) strawberry, d) chokeberry data.

Table 3
The probabilities of the calibration and cross-validation models.

	Berry	Calibration			Cross-validation			LV
		Wilcoxon	Sign Test	Rand *t*-test	Wilcoxon	Sign Test	Rand *t*-test	
Full chromatograms	RB	0.000	0.008	0.005	0.000	0.003	0.005	8
	BC	0.000	0.011	0.005	0.000	0.016	0.005	11
	ST	0.001	0.006	0.005	0.000	0.002	0.005	11
	CHB	0.000	0.003	0.005	0.000	0.002	0.005	8
After VIP variable selection	RB	0.000	0.007	0.005	0.000	0.001	0.005	5
	BC	0.000	0.007	0.005	0.000	0.004	0.005	7
	ST	0.001	0.008	0.005	0.000	0.003	0.005	4
	CHB	0.000	0.002	0.005	0.000	0.002	0.005	5

Note: RB: raspberry; CB: chokeberry; ST: strawberry; BC: blackcurrant; LV: latent variables.

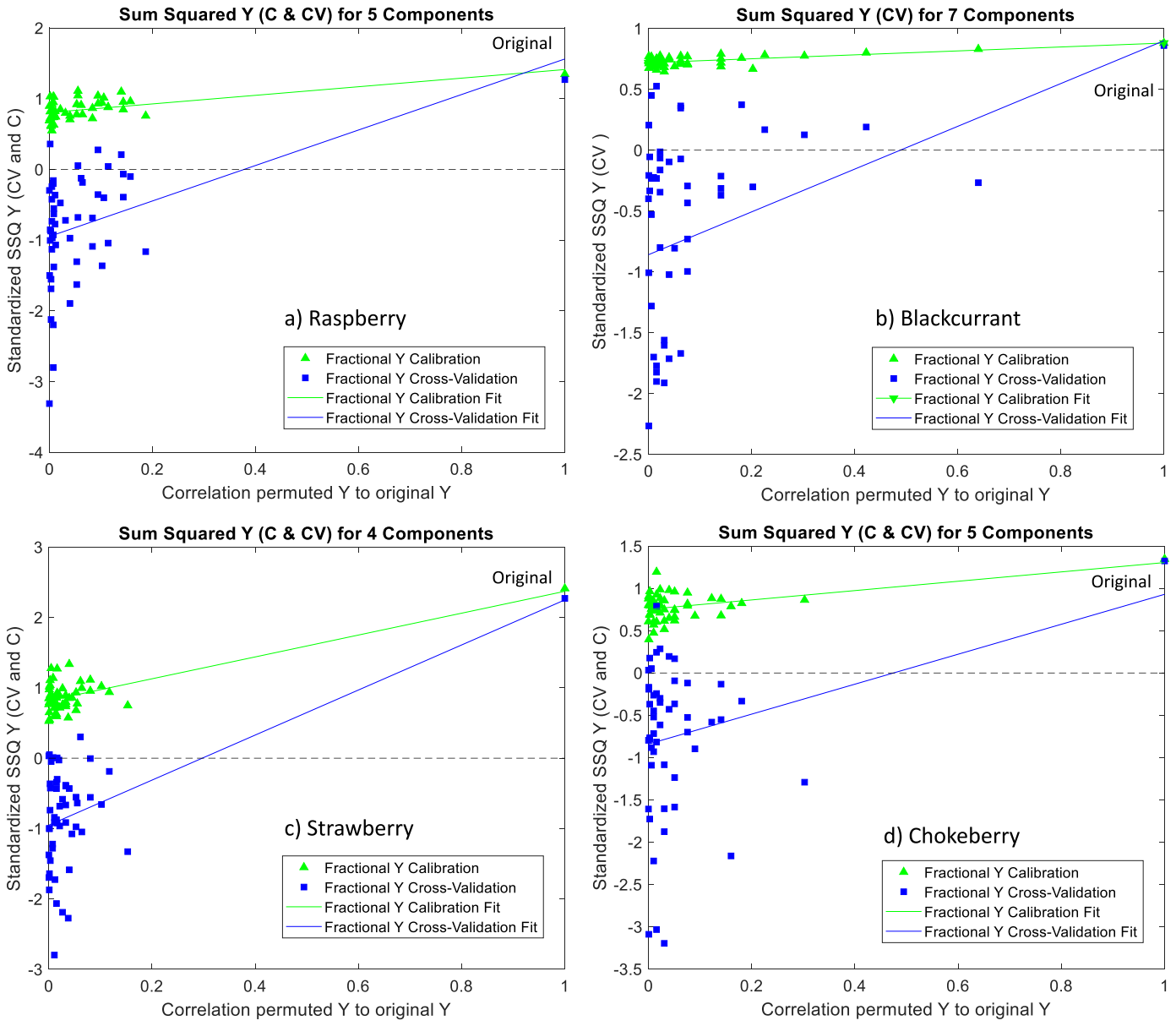

Fig. 6. Fractional y-variance from calibration and cross-validation versus correlation of the permuted y-block to original y-block in a) raspberry, b) blackcurrant, c) strawberry, d) chokeberry seed oils.

CHB samples).

With the aim of reducing the original data set, removing redundant or unnecessary chromatogram regions, and thus increasing the PLS model predictive power, variables characterized by VIP>1 were selected and new PLS models were developed. Chromatograms for selected regions are shown in Fig. 4. The number of variables was reduced by more than 75% (1535 variables) and up of 85% (905 variables) with respect to the original data set (6200 variables). The number of LV is also considerably reduced (from 9.5 ± 1.5 to 5.2 ± 1.1) making the models more stable. The results of the new PLS models are shown in Table 2. For all the samples, except ST, results in cross-validation were improved (mean RMSECV reduction of $26.3 \pm 13.2\%$) compared to those obtained for full chromatograms. R^2 ranges from 0.895 (RMSECV = 1.42 months, ST) to 0.969 (RMSECV = 0.75 months, CHB), while RPD values vary between 3.1 and 5.7. The results in terms of measured versus predicted storage time (months) of the cross validated models are reported in Fig. 5. Furthermore, determination coefficients (R^2) RMSE and BIAS in calibration and cross-validation are shown.

Although there is no statistical basis as to how the threshold equal to two was determined, usually models characterized by RPD values higher than two are considered as excellent. Furthermore, for the quality control field, the following RPD categories have been identified: i) 2.4–3.0 rough screening quality; ii) 3.1–4.9 screening quality; iii) 5.0–6.4 quality control; iv) 6.5–8.0 process control; v) > 8.1 any application (Williams and Norris, 2001). Considering the RPD values achieved in this study, the PLS models could be suitable for predicting the storage time in quality control.

Due to the restricted number of the samples, it was important to evaluate the robustness and significance of the model as a function of the latent variable, and also to avoid overfitting. In particular, the model robustness and significance was verified through a permutation test that shows the probability of the original model (unpermuted) being significantly different from the one built under the same conditions but using random data (permuted model). The probability that the unpermuted model is not significantly different from the one created from randomly shuffling the y-block was evaluated by using three different tests: Pairwise Wilcoxon signed rank test (Wilcoxon), Pairwise signed rank test (Sign Test) and Randomization *t*-test (Rand *t*-test). Table 3 shows the

probabilities in calibration and cross-validation, with values less than 0.05 indicating that the model is significant at the 95% confidence level for the specific number of latent variables. It is possible to confirm that all the developed models are significant with values higher than 98%, both in calibration and validation.

Furthermore, fractional y-variance captured for calibration and cross-validation versus the correlation of the permuted y-block to the original y-block was evaluated as an index of model robustness (Fig. 6, PLS models developed after variable selection). For each permuted y-block, RMSEC and RMSECV were used to calculate fractional sum squared Y captured (SSQ Y) for the calibration (C) and cross-validation (CV). In general, the cross-validated and calibration SSQ Y values should be relatively close to each other but should be less than the results for the unpermuted y-block (right side of the plot), independent of their correlation with the real class values. Consequently, all the models can be considered significant, though the PLS model related to BC shows less robust results in terms of dispersion of SSQ Y values (both C and CV) and the distance between unpermuted and permuted SSQ Y. This agrees with the probability's values reported in Table 3, especially for the Sign Test.

4. Conclusions

The changes in the volatile profile of berry seed oils during one year of storage were correctly evaluated by FGC E-nose combined with chemometric techniques. A clear separation between samples obtained from different suppliers was demonstrated in PCA. Additionally, samples from 0 to 12 months were well distributed, according to the function of storage time. PLS models were built to predict the storage times for each berry type, and the predictive power was improved after pre-processing and selecting the specific region in a chromatogram. The PLS models built in this study possessed excellent predictive power, as characterised by the RPD values that were greater than 2. It indicates the suitability of the PLS models to predict the storage time during quality control. As a rapid and non-targeted instrument, FGC E-nose proved its robust application to analyse the quality differences of berry seed oils during storage. Data processing with chemometrics was able to extract a wider scope of information from the resulting chromatogram, such as visualisation of oil characteristics according to the berry type, supplier origin, and storage prediction. Even though this study using FGC E-nose is limited to qualitative determination, such an approach is valued by the oil industry, which requires analysis of a great number of oils daily, with respect to their quality and authenticity. FGC E-nose is a versatile tool compared to conventional GC, only 100 s of acquisition time are required and no need for solvents.

Funding

Samples purchased using the fund by NATIONAL SCIENCE CENTRE, POLAND, OPUS NCN grant number: 2018/31/B/NZ9/02762.

CRediT authorship contribution statement

Yolanda Victoria Rajagukguk: Conceptualization, Methodology, Investigation, Writing – original draft, Writing – original draft. **Chiara Cevoli:** Conceptualization, Data curation, Formal analysis, Methodology, Validation, Writing – original draft, Writing – original draft. **Ilaria Grigoletto:** Methodology. **Jolanta Tomaszewska-Gras:** Conceptualization, Funding acquisition, Methodology, Writing – review & editing.

Declaration of competing interest

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

Data availability

Data will be made available on request.

References

Aliakbarzadeh, G., Parastar, H., Sereshti, H., 2016. Classification of gas chromatographic fingerprints of saffron using partial least squares discriminant analysis together with different variable selection methods. Chemometr. Intell. Lab. Syst. 158, 165–173. https://doi.org/10.1016/j.chemolab.2016.09.002.

Barbieri, S., Cevoli, C., Bendini, A., Quintanilla-Casas, B., García-González, D.L., Gallina Toschi, T., 2020. Flash gas chromatography in Tandem with chemometrics: a rapid screening tool for quality grades of virgin olive oils. Foods 9, 862. https://doi.org/10.3390/foods9070862.

Cevoli, C., Casadei, E., Valli, E., Fabbri, A., Gallina Toschi, T., Bendini, A., 2022. Storage time of nut spreads using flash gas chromatography E-nose combined with multivariate data analysis. Lwt 159. https://doi.org/10.1016/j.lwt.2022.113217.

Damiani, T., Cavanna, D., Serani, A., Dall'Asta, C., Suman, M., 2020. GC-IMS and FGC-Enose fingerprint as screening tools for revealing extra virgin olive oil blending with soft-refined olive oils: a feasibility study. Microchem. J. 159, 105374 https://doi.org/10.1016/j.microc.2020.105374.

European Commission, 2021. ACN Annual Report. Alert and Cooperation Network.

FAO, 2021. Crops: Raspberries and Strawberries. https://www.fao.org/faostat/en/#data/QV.

Farrés, M., Platikanov, S., Tsakovski, S., Tauler, R., 2015. Comparison of the variable importance in projection (VIP) and of the selectivity ratio (SR) methods for variable selection and interpretation. J. Chemom. 29, 528–536. https://doi.org/10.1002/cem.2736.

Gaca, A., Kludská, E., Hradecký, J., Hajšlová, J., Jeleń, H.H., 2021. Changes in volatile compound profiles in cold-pressed oils obtained from various seeds during accelerated storage. Molecules 26, 1–14. https://doi.org/10.3390/molecules26020285.

Liang, P., Akoh, C.C., Diehl, B.W.K., Jacobsen, C., 2020. Oxidative stability of cod liver oil in the presence of herring roe phospholipids. Food Chem. 310, 125868 https://doi.org/10.1016/j.foodchem.2019.125868.

Majerska, J., Michalska, A., Figiel, A., 2019. Trends in Food Science & Technology A review of new directions in managing fruit and vegetable processing by- products. Trends Food Sci. Technol. 88, 207–219. https://doi.org/10.1016/j.tifs.2019.03.021.

Martysiak-Żurowska, D., Orzołek, M., 2023. The correlation between nutritional and health potential and antioxidant properties of raw edible oils from cultivated and wild plants. Int. J. Food Sci. Technol. 58, 676–685. https://doi.org/10.1111/ijfs.16217.

Marx, Í.M.G., Rodrigues, N., Veloso, A.C.A., Casal, S., Pereira, J.A., Peres, A.M., 2021. Volatile-Olfactory profiles of cv. Arbequina olive oils extracted without/with olive leaves addition and their discrimination using an electronic nose. J. Chem. 2021 https://doi.org/10.1155/2021/5058522.

Mazurek, B., Ryszko, U., Kostrzewa, D., Chmiel, M., Kondracka, M., 2022. Brief characteristics of oxidative stability, fatty acids and metal content in selected berry seed extracts obtained by the SFE technique and used as potential source of nutrients. Food Chem. 367 https://doi.org/10.1016/j.foodchem.2021.130752.

Melucci, D., Bendini, A., Tesini, F., Barbieri, S., Zappi, A., Vichi, S., Conte, L., Gallina Toschi, T., 2016. Rapid direct analysis to discriminate geographic origin of extra virgin olive oils by flash gas chromatography electronic nose and chemometrics. Food Chem. 204, 263–273. https://doi.org/10.1016/j.foodchem.2016.02.131.

Mildner-Szkudlarz, S., Różańska, M., Siger, A., Kowalczewski, P.Ł., Rudzińska, M., 2019. Changes in chemical composition and oxidative stability of cold-pressed oils obtained from by-product roasted berry seeds. Lwt 111, 541–547. https://doi.org/10.1016/j.lwt.2019.05.080.

Modesti, M., Taglieri, I., Bianchi, A., Tonacci, A., Sansone, F., Bellincontro, A., Venturi, F., Sanmartin, C., 2021. E-nose and olfactory assessment: Teamwork or a challenge to the last data? the case of virgin olive oil stability and shelf life. Appl. Sci. 11, 1–20. https://doi.org/10.3390/app11188453.

Mota, M.F.S., Waktola, H.D., Nolvachai, Y., Marriott, P.J., 2021. Gas chromatography – mass spectrometry for characterisation, assessment of quality and authentication of seed and vegetable oils. TrAC - Trends Anal. Chem. 138, 116238 https://doi.org/10.1016/j.trac.2021.116238.

Palagano, R., Valli, E., Cevoli, C., Bendini, A., Gallina Toschi, T., 2021. Compliance with Eu Vs. Extra-Eu Labelled geographical provenance in virgin olive oils: a rapid untargeted chromatographic approach based on volatile compounds. Riv. Ital. delle Sostanze Grasse 98, 312–314. https://doi.org/10.1016/j.lwt.2020.109566.

Przykaza, K., Nikolaichuk, H., Kozub, A., Tomaszewska-Gras, J., Peršurić, Ž., Pavelić, S. K., Fornal, E., 2021. Newly marketed seed oils. What we can learn from the current status of authentication of edible oils. Food Control 130. https://doi.org/10.1016/j.foodcont.2021.108349.

Rajagukguk, Y.V., Islam, M., Siger, A., Fornal, E., Tomaszewska-Gras, J., 2023. Oxidative stability assessment of industrial and laboratory-pressed fresh raspberry seed oil (Rubus idaeus L.) by differential scanning calorimetry. Food Chem. Adv. 2, 100186 https://doi.org/10.1016/j.focha.2023.100186.

Tata, A., Massaro, A., Damiani, T., Piro, R., Dall'Asta, C., Suman, M., 2022. Detection of soft-refined oils in extra virgin olive oil using data fusion approaches for LC-MS, GC-IMS and FGC-Enose techniques: the winning synergy of GC-IMS and FGC-Enose. Food Control 133, 108645. https://doi.org/10.1016/j.foodcont.2021.108645.

Teixeira, G.G., Dias, L.G., Rodrigues, N., Marx, Í.M.G., Veloso, A.C.A., Pereira, J.A., Peres, A.M., 2021. Application of a lab-made electronic nose for extra virgin olive

oils commercial classification according to the perceived fruitiness intensity. Talanta 226, 1–26. https://doi.org/10.1016/j.talanta.2021.122122.

Tian, H., Wu, D., Chen, B., Yuan, H., Yu, H., Lou, X., Chen, C., 2023. Rapid identification and quantification of vegetable oil adulteration in raw milk using a flash gas chromatography electronic nose combined with machine learning. Food Control 150. https://doi.org/10.1016/j.foodcont.2023.109758.

Tomasi, G., Van Den Berg, F., Andersson, C., 2004. Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data. J. Chemom. 18, 231–241. https://doi.org/10.1002/cem.859.

van den Berg, R.A., Hoefsloot, H.C.J., Westerhuis, J.A., Smilde, A.K., van der Werf, M.J., 2006. Centering, scaling, and transformations: improving the biological information content of metabolomics data. BMC Genom. 7, 1–15. https://doi.org/10.1186/1471-2164-7-142.

Williams, P., Norris, K., 2001. Near-infrared Technology in the Agricultural and Food Industries, second ed. American Association of Cereal Chemists, St. Paul, MN.

Xu, L., Yu, X., Li, M., Chen, J., Wang, X., 2018. Monitoring oxidative stability and changes in key volatile compounds in edible oils during ambient storage through HS-SPME/GC–MS. Int. J. Food Prop. 20, S2926–S2938. https://doi.org/10.1080/10942912.2017.1382510.

Zhang, Y., Wu, G., Chang, C., Lv, Y., Lai, W., Zhang, H., Wang, X., Jin, Q., 2020. Determination of origin of commercial flavored rapeseed oil by the pattern of volatile compounds obtained via GC–MS and flash GC electronic nose. Eur. J. Lipid Sci. Technol. 122, 1–7. https://doi.org/10.1002/ejlt.201900332.