Title

Non-destructive method to classify walnut kernel freshness from volatile organic compound (VOC) emissions using gas chromatography-differential mobility spectrometry (GC-DMS) and machine learning analysis

Journal

Applied Food Research, 3(2)

ISSN

2772-5022

Authors

Chakraborty, Pranay
Borras, Eva
Rajapakse, Maneeshin Y
et al.

Publication Date

2023-12-01

DOI

10.1016/j.afres.2023.100308

Published in final edited form as:
Appl Food Res. 2023 December ; 3(2): . doi:10.1016/j.afres.2023.100308.

Non-destructive method to classify walnut kernel freshness from volatile organic compound (VOC) emissions using gas chromatography-differential mobility spectrometry (GC-DMS) and machine learning analysis

Pranay Chakraborty[1,2,§], **Eva Borras**[1,2], **Maneeshin Y. Rajapakse**[1,2], **Mitchell M. McCartney**[1,2,3], **Matthew Bustamante**[4], **Elizabeth J. Mitcham**[4], **Cristina E. Davis**[1,2,3,*]

[1]Department of Mechanical and Aerospace Engineering, University of California Davis, Davis, CA, USA

[2]UC Davis Lung Center, One Shields Avenue, Davis, CA, USA

[3]VA Northern California Health Care System, 10535 Hospital Way, Mather, CA, USA

[4]Department of Plant Sciences, University of California, Davis, Davis, CA, USA

Abstract

Analysis of volatile organic compounds (VOCs) can be an effective strategy to inspect the quality of horticultural commodities and following their degradation. In this work, we report that VOCs emitted by walnuts can be studied using gas chromatography-differential mobility spectrometry (GC-DMS), and those GC-DMS data can be analyzed to predict the rancidity of walnuts, i.e., classify walnuts into grades of freshness. Walnut kernels were assigned a class n depending on their level of freshness as determined by a peroxide assay. VOC samples were analyzed using GC-DMS. From these VOC data, a partial least square regression (PLSR) model provided a freshness prediction value m, which corresponded to the rancid class n when $m = n \pm 0.5$. The PLSR model had an accuracy of 80% to predict walnut grade and demonstrated a minimal root mean squared error of 0.42 for the m response variables (representative of walnut grade) with the GC-DMS data. We also conducted gas chromatography-mass spectrometry (GC-MS) experiments to identify volatiles that emerged or were enhanced with more rancid walnuts. The findings of the GC-MS study of walnut VOCs align excellently with the GC-DMS study. Based on our results,

---

[*]Corresponding author: cedavis@ucdavis.edu.
[§]Presently at Southern Arkansas University, Magnolia, AR 71753, USA

Competing Interests: The authors declare no competing financial interest.

Data and software availability: The software code for AIMS is available on GitHub for non-commercial use. Please refer to Professor Cristina Davis' webpage for more information. The Regents of the University of California own the copyrights to the software. Future published scientific manuscripts or reports using this software and/or hardware designs must cite this the original publication (DOI: 10.1007/s12127-016-0200-9), previous configuration reports (DOIs: 10.1016/j.chemolab.2020.104085, 10.1021/acs.analchem.9b01428, 10.1039/C8AY00846A and 10.1039/d2ay00723a, and 10.1039/D2AY00723A). Experimental data requests should be addressed to the corresponding author.

we conclude that a GC-DMS device deployed with a pre-trained machine learning model can be a very effective device for classifying walnut grades in the industry.

Keywords

machine learning; differential mobility spectrometry; mass spectrometry; peroxide value; rancidity

1. Introduction

Walnuts (*Juglans regia* L.) are among the top nut crop produced globally [1, 2]. Their kernels are known to have antioxidant qualities due to high levels of *n*-3 polyunsaturated fatty acids (PUFA). However, these nutritionally beneficial fats are prone to oxidation and hydrolysis, generating byproducts that cause off-flavors, rancidity, and a less desirable product. Oxidation of walnut phenols darkens the appearance of kernels, impacting marketability. Walnut researchers have demonstrated methods that reduce the likelihood of oxidation, such as drying walnuts immediately after harvest [3] or optimizing the relative humidity and temperature conditions during storage [4].

The walnut industry determines the degree of rancidity through analysis of free fatty acids and peroxide values from extracted walnut oil. These methods require sample preparation steps such as solvent extraction that are done in a laboratory and delay results. A more rapid technique to quantify rancidity at point-of-sampling would benefit the walnut industry by allowing growers or distributors to quickly sort walnuts in terms of a grade of product, e.g., a premium “freshest” product versus spoiled walnuts.

VOC-based studies have been very successful 0.4in determining the food quality or detecting the onset of the rancidification of foods, e.g., by detecting meat spoilage [5], diseases in potato tubers [6-8], and fungal disease in mango [9]. Lu et al. [10] adopted a gas chromatography-tandem mass spectrometry (GC-MS/MS) approach to study the contamination of fruits and vegetables with a wide range of pesticides. In a recent study, Grilo et al. conducted GC-MS experiments to evaluate potential oxidation markers of walnuts during a 28-week storage period [11]. In a different study, Grilo et al. [12] adopted the rancimat method [13] in differentiating oxidative stability of walnut kernels to inspect walnut deterioration.

Sophisticated VOC analysis instrumentations like gas chromatography and mass spectrometry, laser spectrometry, and differential mobility spectrometry (DMS) are being utilized with multivariate data processing tools for the quality check of a wide range of food products, including animal and plant products, beverages, and honey [14-17]. Among the existing technologies to analyze VOCs, gas chromatography-differential mobility spectrometry (GC-DMS) is gaining significant attention for several reasons. In particular, the functionality of DMS sensors at ambient temperatures and pressures and continuous miniaturization makes the GC-DMS an excellent candidate for lightweight and portable VOC analysis sensors. Furthermore, VOC samples can be collected from walnut kernels nondestructively and does not require sample preparation procedures.

In a GC-DMS device, a GC column precedes the DMS to allow the pre-separation of volatile compounds before orthogonal separation and detections in DMS. In a DMS, VOCs are first ionized, then ions are distinguished by the variation of ion mobilities at a high and low electric field. DMS is a variation of ion mobility spectrometry (IMS) with two planner electrodes that establish an asymmetric waveform with an electric field of ~30 KV/cm at MHz frequency. A detailed comparison of conventional IMS and DMS, along with their pros and cons, can be found in the existing literature [18-21]. Because of the field dependence of ion mobility, ions tend to be dragged toward one of the two planner electrodes. Such drift is compensated by an additional constant voltage (which is known as compensation voltage, CV) to allow ions to pass through the electrode gap and reach the detector. Depending on the nature of ions, they are detected at different CVs. The details of the working principles of the DMS and their miniaturization are well reported in the literature [22-26].

A hyphenated GC and DMS works in series to result in orthogonal separation of VOCs in a 3-dimensional plot. Specifically, a GC-DMS device generates 3-dimensional biomarker plots with axes representing retention time, compensation voltage, and ion signal intensity. The easy sample preparation, simple operation techniques, and less constricted working conditions can make GC-DMS a viable technology for devising sensors that can be used to analyze walnut VOCs for industry applications.

In this work, we demonstrate volatile compounds emitted by walnuts can be analyzed with a GC-DMS device, and machine learning models can predict the grade of walnuts from GC-DMS data. The GC-DMS-based walnut VOC study has been verified using gold standard GC-MS data and peroxide values of different grades of walnuts. Our work will be a possible benchmark for developing a simpler machine learning-assisted walnut VOC analysis device that can be applied for industrial uses.

2. Materials and methods

2.1. Walnut samples

For our study, we used commercial walnuts that were kept as kernels for a certain number of months. The kernel-storage duration was used as a walnut rancidity index. Our rationale for grading the rancidity index based on the storage time is that it is reasonable to assume that rancidity should increase over time. All walnuts were stored with shells for five months prior to storing them as kernels. All walnuts were stored with shells for five months before storing them as kernels at room temperature for four different durations to create four grades of walnuts. The resulting four grades of walnuts are immediately stored at a very low temperature (−40°C) to minimize additional degradation. The details of all the walnuts of different grades are provided in Table 1.

2.2. Headspace (HS) GC-DMS data collection and analysis

We collected the HS gas chromatography-differential mobility spectrometry (GC-DMS) data for different grades of walnuts using a Sionex device (version 2.4.3, Sionex Corp., Bedford, MA). The GC-DMS device consists of a Carbopack X / Carbopack B sorbent preconcentrator trap connected to a DB-XLB 10 m × 0.25 mm × 1.0 µm GC column interfaced to a DMS sensor.

The details of different grades of samples are described in section 2.1. Initially, we placed 30 grams of walnut samples in a 500 mL borosilicate glass jar and kept them for 30 mins at room temperature to allow walnut VOCs to equilibrate with the headspace. After equilibration of the walnut samples, the jar was connected to the inlet of the GC-DMS device using Teflon tubing and Swagelok fittings to establish a leak-free connection. A schematic of the experimental setup is shown (Figure 1).

Once the jar was connected to the GC-DMS device, the sample pump was run for a sampling time of 120 s to trap the VOCs into the device preconcentrator trap. The preconcentrator trap of the GC-DMS device performs the precontraction step by trapping and concentrating VOCs from the walnut jar onto sorbent material to facilitate improved limits of detections. Then the sample trap temperature (initially at 40 °C) was suddenly raised to 300 °C and kept at 300 °C for 180 s. The high temperature of the trap desorbs (releases) the entrapped VOCs, which then pass through the GC column for initial separation. The GC temperature was then ramped from 40 °C (initial temperature) to 140 °C within 600 s and kept at 140 °C for 670 s.

Once the VOCs travel through the GC column, they enter the DMS cell, and the GC eluent is mixed with the DMS carrier gas (filtered air) and passed through a 5 mCi ^{63}Ni ionization source. Resulting ions travel through a varying electric field. The filter gap (distance between the two parallel electrodes) and the maximum field strength of the DMS were 500 µm and 110 Townsends, respectively. The compensation voltage (CV) of the DMS was scanned from −30 V to 10 V with a step size of 0.4 V. We performed the CV scan with a step duration, scan duration, and step settle time of 10 ms, 1000 ms, and 3 ms, respectively. The radiofrequency (RF) voltage (also known as separation voltage) was set to 700 V. The sensor temperature was set to 80 °C. After the completion of DMS CV scans for each GC run, we receive 3-dimensional GC-DMS plots of the VOCs of walnuts which were then processed and analyzed. We performed the process mentioned above five times to collect five sets of GC-DMS data for each grade of walnuts. It is worth noting that we did not reuse the same walnuts for replicate data collection.

We conducted GC-DMS data preprocessing with a lab-custom AnalyzeIMS (AIMS) software [27-31]. Specifically, we performed noise reduction of the data through the Savitzky-Golay filter. We performed baseline correction to separate the actual signals from interference effects or remove background effects. For baseline correction of the experimental data, we adopted asymmetric least squares smoothing algorithm. We performed principal component analysis (PCA) and partial least square regression (PLSR) analysis to study the feasibility of machine learning models to differentiate the VOCs of different grades of walnuts. It is worth noting that PCA creates variables without considering the response variable (grades in our study) [32]. On the other hand, partial least square regression (PLSR) takes the response variable while constructing the new predictor variables [32]. The root mean square error of the predicted values from the PLSR analysis is calculated as

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{1}$$

in which MSE is the mean squared error, y_i and $\hat{y}_i$ are the actual and predicted values, respectively, and n is the number of samples. We adopted leave-one-out cross validation approach to check the robustness of our model. In particular, we split the dataset keeping one observation for testing and the rest of the data for training a model. We repeat the process for n times (where n is the total number of observations in the data set), and in this way, we get a less biased measure of test MSE compared to using a single test set. Since PLSR is a regression analysis, we use a tolerance limit to assign a specific grade. More specifically, if the predicted value lies within $m \pm 0.5$, we assign the walnut grade to be m.

2.3. HS-SPME GC-MS data collection and analysis

GC-MS data allows the identification of compounds related to rancidity. These specific compounds allow a better understanding of the grades defined. For that, VOCs from walnuts were collected using modified methods based on literature [11, 33] that defines the main volatiles related to walnut oxidation. Specifically, we extracted volatiles from 2g of walnut kernels in 20 mL glass vials with a PTFE/silicone septum. After 10 mins of equilibration at 40 °C, a DVB/CAR/PDMS solid-phase microextraction (SPME) fiber (Supelco, Bellafonte, PA, USA, MA) was exposed for 40 min to the headspace of the sample. VOCs were analyzed on a 7890B gas chromatograph (GC, Agilent Technologies) coupled to a 5977A quadrupole mass spectrometer (MS). Compounds were separated using a DB-5MS capillary column (30 m × 250 μm i.d., film thickness 0.25 μm,) and helium at a constant flow rate of 1 mL/min. Oven temperature started at 40 °C for 5 min, ramped at 5 °C/min to 120 °C, and then ramped at 15 °C/min to the final temperature of 280 °C, with a final holding time of 10 min. The injector temperature was 260 °C working in a split-less mode. VOCs were then measured with the MS operating in the electron impact mode with a transmission line temperature of 300 °C. Chromatograms were collected scanning from m/z 50 to 500.

Blank injections (no fiber) and blank SPMEs (using clean fiber) were injected throughout the sample sequence. Each sample of kernels was analyzed in triplicates, i.e., 3 SPME fibers per walnut sample. Data was analyzed using Agilent MassHunter Qualitative Analysis. The volatile compounds were tentatively identified using the NIST 2020 Mass Spectral Library. To extract the features of these samples, deconvolution and alignment of the GC–MS data files were performed using recursive feature extraction methodology on Profinder (Version B.08.00, Agilent Technologies Inc.) and Mass Profiler Professional (MPP, V13.0). This process provides a peak table with samples in columns and variables or features in rows, while each entry in the table is peak areas or intensities corresponding to a sample and variable combination. The peak table was then cleaned by removing features that are missing two of the three replicates and appear in blank samples.

2.4. Peroxide value (PV) measurement of walnut oils

The peroxide value (PV) is a standard chemical index for portraying the oxidative deterioration of oils. In the current work, we measured the PVs of walnut oils to correlate the PVs and walnut rancidity. Initially, ten walnuts were chopped into smaller pieces so that they could fit into the stainless steel test-cylinder-outfit of a bench-top hydraulic lab press (model #3925, Carver Inc., Wabash, IN). Oil (14 mL) was extracted and placed into a plastic 15mL freezer safe tube. The air in the tube was nitrogen flushed and capped before being stored in a −80°C freezer. For peroxide value (PV) analysis, frozen walnut oil was thawed fully, and 5 ml of oil was analyzed for PV by titrating with sodium thiosulfate following the American Oil Chemists' Society (American Oil Chemists' Society, 2017) Official Method (Cd 8-53).

3. Results and Discussion

In our work, we measure VOCs emitted by walnut kernels of varying grades of rancidity/ freshness. VOCs were measured by GC-DMS and the resulting data were used to train and validate a PLSR model to predict walnut grade. Then, GC-MS analysis was conducted to confirm the GC-DMS results and to identify walnut compounds related to rancidity.

3.1. Gas chromatography/ differential mobility spectrometry (GC-DMS) plots of different grades of walnuts

In Figure 2, we report GC-DMS plots of the VOCs of the four different grades of walnuts studied. It is worth noting that grade 1 and grade 4 walnuts are the freshest and most rancid walnuts, respectively, and the other two grades (grade 2 and grade 3) have intermediate rancidity. We observe the separation of VOC ions in both the retention time (RT) and compensation voltage (CV) directions. With bare eyes, we see that the intensity of several ions (represented by peaks in GC-DMS plots) increases as rancidity increases (grade 1 to grade 4). To have a quantitative idea we plot the intensity changes of peak P1 and P2 as a function of rancidity index (RI) (Figure 3). Clearly, peak intensities for P1 and P2 increase as the rancidity increases.

3.2. Machine learning assisted classification of walnuts using GC-DMS data

While it is tempting to use several easily detectable by visual inspection peak intensities like P1 and P2 (Figure 2) as the only measure of walnut rancidity, this strategy can lead us to miss all the less visible but potentially relevant peaks. As such, we perform principal component analysis (PCA), a dimensionality reduction technique, to examine whether GC-DMS signals of walnut volatiles can be clustered based on their grades (Figure 4). From Figure 4a, we observe that grade 1 and grade 2 walnuts tend to separate well from each other though they have slight overlaps. However, the PCs of grade 3 and grade 4 overlap significantly and do not form discernible clusters. One plausible explanation of this finding is that grade 3 and grade 4 walnut were similarly rancid, resulting in similar GC-DMS profiles. Another reason may be that all the walnuts of a particular grade may not have become equally rancid. To check our arguments, we analyzed the VOCs of different grades of walnuts through GC-MS experiments and report the peroxide values (PV) of the considered grades of walnuts (section 3.3).

It is worth noting that we expect excellent clustering through PCA analysis if we use walnuts with very distinct (grade 1 and grade 4 for our study) rancidity. Indeed, grade 1 (referred to as freshest) and grade 4 (referred to as most rancid) separate along PC1 (Figure 4b), and this works as a great reference for developing a machine learning model that is able to differentiate rancid and fresh walnuts based on their VOCs.

We now perform PLSR analysis of the GC-DMS data of the VOCs of different grades of walnuts to correlate the GC-DMS data and the walnut grades (Figure 5). We use nineteen sets of data for training the PLSR model and predict the class of another data point with the developed model. Specifically, we developed 20 different models with a leave-one-out approach, and we observed that our machine learning models could predict the walnut grades accurately for 80% of cases. It is noteworthy that no modeled data have been used for prediction purposes. We compare the actual class of walnuts and predicted class response variable of walnuts for the PLSR approach in Figure 5, and the root mean squared error of prediction was 0.42. Based on this tolerance-based predicted class (discussed in the methodology section), we observe that our machine learning model can predict the walnut grades accurately for 80 % of cases. We want to point out that even though our data set appears to be small in this pilot study, the classification accuracy of our machine learning model is reasonably good. One important observation of the PLSR analysis is that it aligns well with PCA analysis findings (Figure 4). In particular, since grade 3 and grade 4 walnut VOCs are similar, we do not observe excellent grade predictions for these classes. As such, we recommend that we should always use sufficiently distinct standard rancidity to develop machine learning models for walnut classification purposes.

It is noteworthy that the GC-DMS data collected as is did not provide identification information for the chemicals detected. Even though a GC-DMS device can identify specific chemicals and their concentration based on a pre-built library, that is beyond the scope of our current study. Instead, we used GC-MS for identifiable chemical information on walnut VOCs, and the findings are reported in the following section, which confirms GC-DMS experiment findings.

3.3. Verification of rancidity with Peroxide value measurement and gas chromatography-mass spectrometry (GC-MS) study

The oxidative rancidity of walnut, related to the evolution of nut quality during storage time, is commonly measured by different parameters, such as peroxide value, fatty acid content, sensory analysis, or volatile content. These volatile compounds are mainly formed by the primary oxidation of fatty acids and their further break down into carbonyl compounds (alcohols, aldehydes, and ketones). Those smaller compounds are known to impact the flavor of the nuts [33, 34], and can be detected and identified using HS-SPME-GC-MS.

3.3.1. Evolution of peroxide value (PV) with walnut rancidity increment—We observed that the PVs of walnut oils increases as the rancidity increases, in general (Figure 6). While all grades had some PV value overlap with neighbors, there was more overlap between grades 3 and 4. On the other hand, grade 1 and grade 4 walnut oil have very distinct PVs. Yet, per a one-way ANOVA (analysis of variance), all four grades had significantly different mean peroxide values ($p=3.5 \times 10^{-20}$). Therefore, we can infer that the reference walnut grade samples for GC-DMS experiments and subsequent machine learning model building can be thoroughly investigated based on their PV values to get excellent predictions.

3.3.2. HS-SPME GC-MS study to predict walnut grade—

An initial data visualization of the GC-MS data showed clear chromatographic differences among the VOC profiles of several grades of walnuts (Figure 7 and Figure 8). For example, the peak size of hexanal, the chemical with the highest intensity, increases considerably as the walnut grade changes from 1 to 4. In a similar way, we can identify 1-pentanol, acetic acid, or 1-octen-3-ol; all reaching higher intensities at grade 4. There are several other compounds, which show no peaks in grade 1 while appear throughout the next grades, which are explained later in this section.

After the initial raw data check, we deconvolute and align the samples obtaining a raw dataset with 164 features (compounds). The raw data is then filtered, and the number of features reduces to 126, corresponding to potential marker compounds of the walnut grade. An exploratory PCA model of the filtered data (Figure 9a) shows good differentiation of walnut grades. The fresher nuts (grade 1) show major differences from the rest of grades (PC1 explained 72.54 %). The second principal component, PC2 (which explains 10.94 % variance), explains the differences between grades 2, 3 and 4, being grade 4, with higher rancidity, the one more differentiated of the rest. Grades 2 and 3 present more similarities between them.

The tolerance-based PLSR analysis (Figure 9b) of the GC-MS data results in a prediction accuracy of 66 %, and this accuracy is significantly lower than the GC-DMS based prediction. Three primary plausible reasons for such discrepancy can be: a) small sample sizes, b) different walnut samples were used for GC-MS and GC-DMS, and c) close rancidity among the grades. We summarize a group of 59 significant features in Figure 8. These major features explained the walnut grades difference presented in the multivariate models (PCA and PLSR analysis in Figure 9) and are considered potential markers of the walnut grades.

We identified compounds with scores higher than 65% by comparison of our obtained mass spectra to the NIST 2020 mass spectral library. A heatmap shows the differences in intensities for each individual compound through the grades (Figure 8). We can also find compounds that are not detected in some of the cases, for example, alkanes like 2,5,9-trimethyl-decane and 2,6,10,14-tetramethyl-heptadecane (compounds # 13 and 14) were only detected in grade 4 nuts, while compounds # 40 to 46 were only detected in higher grades, but not in walnuts of grade 1. Some of these compounds have been reported in the literature as compounds associated with the oxidation process of walnuts through time, such as 2-hexenal and (E)-2-pentenal [11]. Differently, only a few compounds like hexanoic acid and methyl 1-methylcyclopropanyl ketone were with higher intensities only in grade 1, when the walnuts were fresher, and the signal reduced afterward. In the case of toluene, it also appeared with a high signal in grade 1 to reduce its amount in grade 2; however, it increased again to its highest intensity in grade 3. Although most of the compounds presented higher intensities at larger storage times, we find a lot of variability within intensities through grades 2, 3, and 4. This confirms the good ability to discriminate samples from grade 1 compared to the rest of the grades shown in PC1 (Figure 9a). Also, several of the volatiles detected (compounds # 11 to 29) showed clear higher intensities when they reached grade 4. This matches with the enhanced ability to discriminate grade 4 from grades 2 and 3 when observed PC2 (Figure 9a). Some of these compounds can be found in literature as markers of oxidation process in walnuts, such as 2-pentyl-furan or (E)-2-octenal [11].

Based on the GC-MS data for different grades of walnuts, we see that the GC-DMS provides a comparatively better walnut grade prediction accuracy than GC-MS. Therefore, a machine learning model-integrated GC-DMS device can be an excellent fast, lightweight, and field-deployable option to check the quality of walnuts or similar products.

4. Conclusions

In this work, we conducted GC-DMS experiments for detecting VOCs from different grades of walnuts and found that GC-DMS data can be used to develop a machine learning model to predict the grades of walnuts in storage with an accuracy of 80%. We observed that GC-DMS provides better accuracy in predicting the walnut grades compared to GC-MS (66%). The peroxide value evolution of different grades of walnut oils affirms our findings from the GC-DMS study. Therefore, a machine learning model-assisted GC-DMS device will be an effective tool for industrial purposes to differentiate grades of walnuts or similar products based on their storage time or rancidity.

Acknowledgments

This work was primarily supported by the California Walnut Board (CWB) Grades & Standards Committee, awards A21-4128-001 [EJM], A19-4599-001 [EJM] and A23-2583 [CED]. This work was also partially supported by: NIH NCATS 1U18TR003795-01, 4U18TR003795-02, 1U01TR004083-01 [CED] and UL1 TR001860 [CED]; NIH award UG3-OD023365 [CED]; NIH award 1P30ES023513-01A1 [CED]; the Department of Veterans Affairs award I01 BX004965-01A1 [CED]; and the University of California Tobacco-Related Disease Research Program award T31IR1614 [CED]. The contents of this manuscript are solely the responsibility of the authors and do not necessarily represent the official views of the funding agencies.

References:

1. FAOSTAT, in Food and Agriculture Organiation of the United Nations. 2022.
2. Anttalainen O., et al. , Possible strategy to use differential mobility spectrometry in real time applications. International Journal for Ion Mobility Spectrometry, 2020. 23(1): p. 1–8.
3. Zhou X., et al. , Comparative analyses of three dehydration methods on drying characteristics and oil quality of in-shell walnuts. Drying Technology, 2018. 36(4): p. 477–490.
4. Adkison C., et al. , Optimizing Walnut Storage Conditions: Effects of Relative Humidity, Temperature, and Shelling on Quality after Storage. HortScience horts, 2021. 56(10): p. 1244–1250.
5. Mayr D., et al. , Rapid detection of meat spoilage by measuring volatile organic compounds by using proton transfer reaction mass spectrometry. Applied and environmental microbiology, 2003. 69(8): p. 4697–4705. [PubMed: 12902260]
6. Kushalappa A., et al. , Volatile fingerprinting (SPME-GC-FID) to detect and discriminate diseases of potato tubers. Plant Disease, 2002. 86(2): p. 131–137. [PubMed: 30823309]
7. Lui L., et al. , Discrimination of three fungal diseases of potato tubers based on volatile metabolic profiles developed using GC/MS. Potato Research, 2005. 48(1): p. 85–96.

8. Rutolo M., et al. , Detection of potato storage disease via gas analysis: a pilot study using field asymmetric ion mobility spectrometry. Sensors, 2014. 14(9): p. 15939–15952. [PubMed: 25171118]
9. Moalemiyan M, Vikram A, and Kushalappa A, Detection and discrimination of two fungal diseases of mango (cv. Keitt) fruits based on volatile metabolite profiles using GC/MS. Postharvest Biology and Technology, 2007. 45(1): p. 117–125.
10. Lu D., et al. , A fast and easy GC-MS/MS method for simultaneous analysis of 73 pesticide residues in vegetables and fruits. Analytical Methods, 2013. 5(7): p. 1721–1732.
11. Grilo FS and Wang SC, Walnut (Juglans regia L.) volatile compounds indicate kernel and oil oxidation. Foods, 2021. 10(2): p. 329. [PubMed: 33557108]
12. Simoes Grilo F, Srisaard Y, and Wang SC, Prediction of walnut deterioration using kernel oxidative stability. Foods, 2020. 9(9): p. 1207. [PubMed: 32882834]
13. Liang C and Schwarzer K, Comparison of four accelerated stability methods for lard and tallow with and without antioxidants. Journal of the American Oil Chemists' Society, 1998. 75(10): p. 1441–1443.
14. Pastor K, Ačanski M, and Vujić D, Gas chromatography in food authentication, in Gas Chromatography-Derivatization, Sample Preparation, Application. 2019, IntechOpen. p. 109.
15. Solis-Solis HM, et al. , Discrimination of eight varieties of apricot (Prunus armeniaca) by electronic nose, LLE and SPME using GC–MS and multivariate analysis. Sensors and Actuators B: Chemical, 2007. 125(2): p. 415–421.
16. Friberg J., Volatile organic compounds analyzed by gas chromatography-deep ultraviolet spectroscopy. Acupuncture and Related Therapies, 2014. 2(1): p. 25–28.
17. Vallarino JG, et al., Acquisition of volatile compounds by gas chromatography–mass spectrometry (GC-MS), in Plant metabolomics. 2018, Springer. p. 225–239.
18. Ayodeji I., et al., Rapid pre-filtering of amphetamine and derivatives by direct analysis in real time (DART)-differential mobility spectrometry (DMS). 2017. 9(34): p. 5044–5051.
19. Campbell JL, Le Blanc JY, and Kibbey RGJB, Differential mobility spectrometry: a valuable technology for analyzing challenging biological samples. 2015. 7(7): p. 853–856.
20. Schneider BB, et al., Planar differential mobility spectrometer as a pre-filter for atmospheric pressure ionization mass spectrometry. 2010. 298(1-3): p. 45–54.
21. Schneider BB, et al., Maximizing ion transmission in differential mobility spectrometry. 2017. 28(10): p. 2151–2159.
22. Krylov E, Nazarov E, and Miller R, Differential mobility spectrometer: Model of operation. International Journal of Mass Spectrometry, 2007. 266(1-3): p. 76–85.
23. Schneider BB, et al. , Planar differential mobility spectrometer as a pre-filter for atmospheric pressure ionization mass spectrometry. International journal of mass spectrometry, 2010. 298(1-3): p. 45–54. [PubMed: 21278836]
24. Haack A., et al. , A first principle model of differential ion mobility: The effect of ion-solvent clustering. Journal of The American Society for Mass Spectrometry, 2019. 30(12): p. 2711–2725. [PubMed: 31755046]
25. Buryakov I., et al. , A new method of separation of multi-atomic ions by mobility at atmospheric pressure using a high-frequency amplitude-asymmetric strong electric field. International Journal of Mass Spectrometry and Ion Processes, 1993. 128(3): p. 143–148.
26. Nazarov EG, et al. , Miniature differential mobility spectrometry using atmospheric pressure photoionization. Analytical Chemistry, 2006. 78(13): p. 4553–4563. [PubMed: 16808465]
27. Peirano DJ, Pasamontes A, and Davis CE, Supervised semi-automated data analysis software for gas chromatography/differential mobility spectrometry (GC/DMS) metabolomics applications. International Journal for Ion Mobility Spectrometry, 2016. 19(2): p. 155–166. [PubMed: 27799845]
28. Yeap D., et al. , Peak detection and random forests classification software for gas chromatography/differential mobility spectrometry (GC/DMS) data. Chemometrics and Intelligent Laboratory Systems, 2020. 203: p. 104085. [PubMed: 32801407]
29. Yeap D., et al. , Machine vision methods, natural language processing, and machine learning algorithms for automated dispersion plot analysis and chemical identification from complex mixtures. Analytical chemistry, 2019. 91(16): p. 10509–10517. [PubMed: 31310101]

30. Rajapakse MY, et al. , Automated chemical identification and library building using dispersion plots for differential mobility spectrometry. Analytical Methods, 2018. 10(35): p. 4339–4349. [PubMed: 30984293]
31. Chakraborty P., et al. , Machine learning and signal processing assisted differential mobility spectrometry (DMS) data analysis for chemical identification. Analytical Methods, 2022.
32. Murphy KP, Machine learning: a probabilistic perspective. 2012: MIT press.
33. Hao J., et al. , HS-SPME GC–MS characterization of volatiles in processed walnuts and their oxidative stability. Journal of food science and technology, 2020. 57(7): p. 2693–2704. [PubMed: 32549619]
34. Shahidi F and John J, Oxidative rancidity in nuts, in Improving the safety and quality of nuts. 2013, Elsevier. p. 198–229.

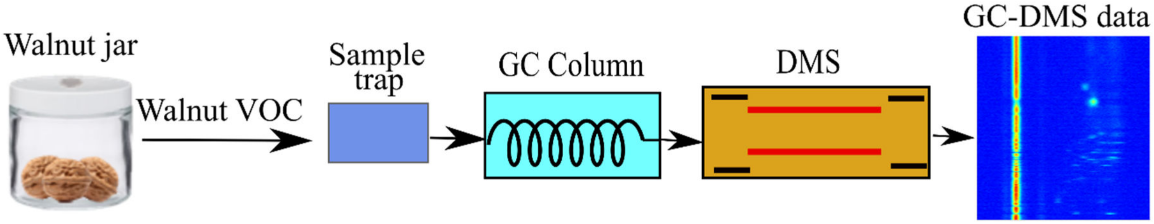

Figure 1:

Schematic of a GC-DMS experimental setup to extract the 3-dimensional biomarker plot of walnut VOCs.

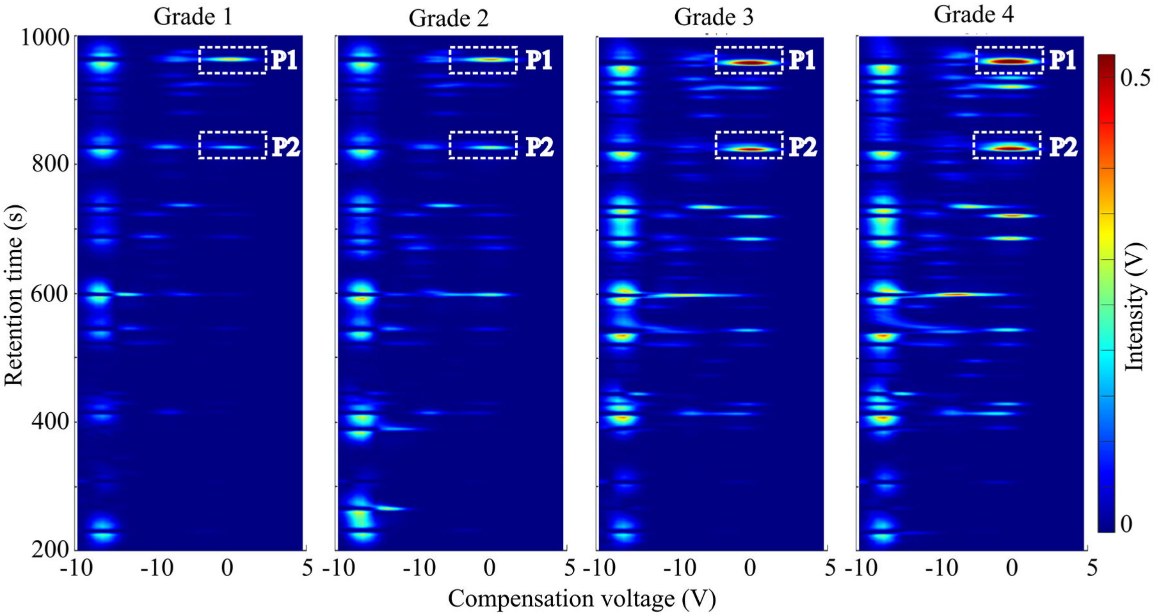

Figure 2:

Gas chromatography/differential mobility spectrometry (GC-DMS) plots of the volatile organic compounds emitted from four different grades of walnuts. Grade 1 was the freshest walnuts, and Grade 4 was the most rancid walnuts. Two boxes titled P1 and P2 are guides for eyes to follow the VOCs concentration changes as walnut grade changes.

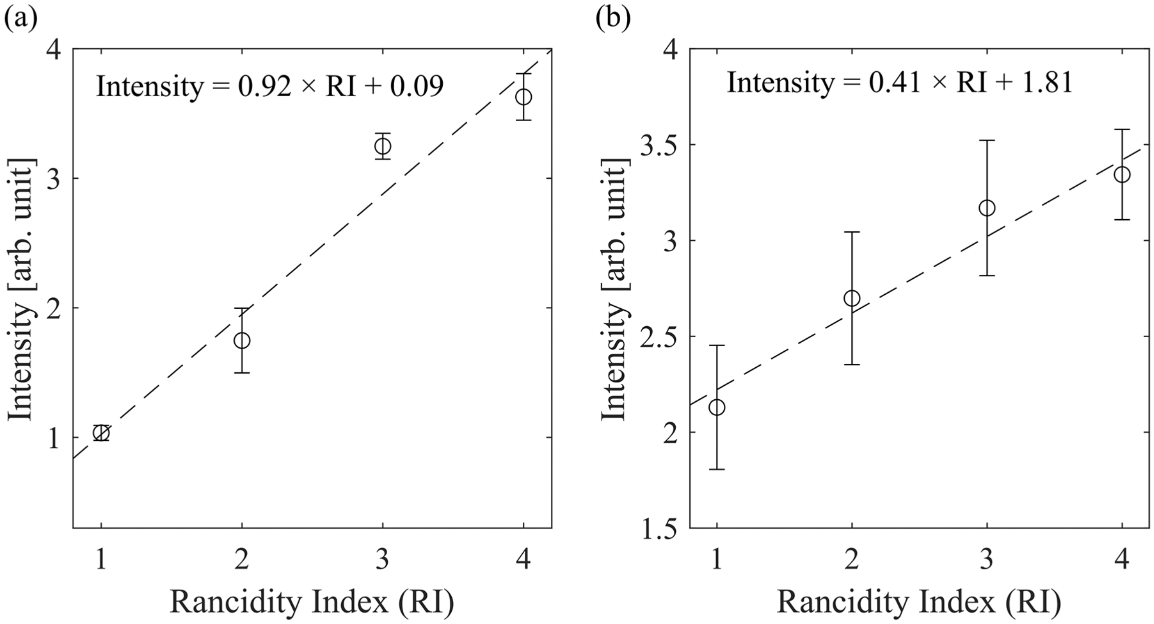

Figure 3:

Intensity change of (a) peak 1 (P1) as a function of rancidity index and (b) peak 2 (P2) as a function of rancidity index. P1 and P2 represent the two ion peaks, as demonstrated in Figure 2. Rancidity indexes of 1, 2, 3, and 4 represent the grade 1 (freshest), grade 2, grade 3, and grade 4 (most rancid) walnuts, respectively.

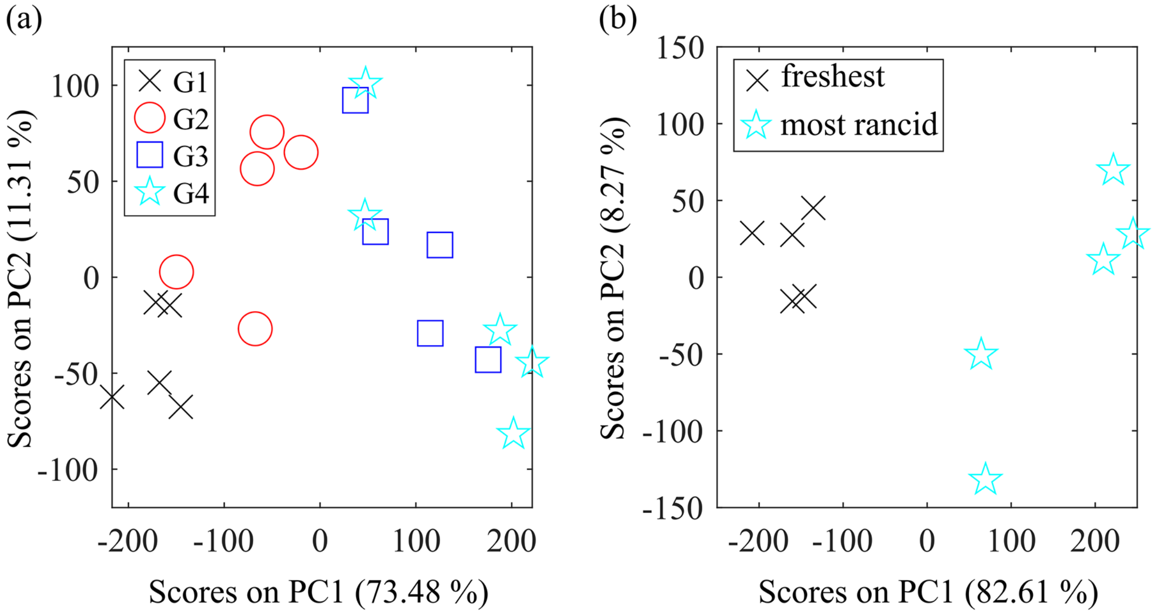

Figure 4:

(a) Scores of the first and second principal components (PC) resulting from the principal components analysis of the GC-DMS data of the VOCs of different grades of walnuts. G1, G2, G3, and G4 represent the grade 1, grade 2, grade 3, and grade 4 walnuts, respectively. PC1 and PC2 explained 73.48 % and 11.31 % of data variance, respectively. (b) Scores of the first and second principal components resulting from the principal components analysis of the GC-DMS data of the VOCs of rancid and nonrancid walnuts. G1 and G4 grades are marked as the freshest and most rancid walnuts. PC1 and PC2 explained 82.61 % and 8.27 % of data variance, respectively.

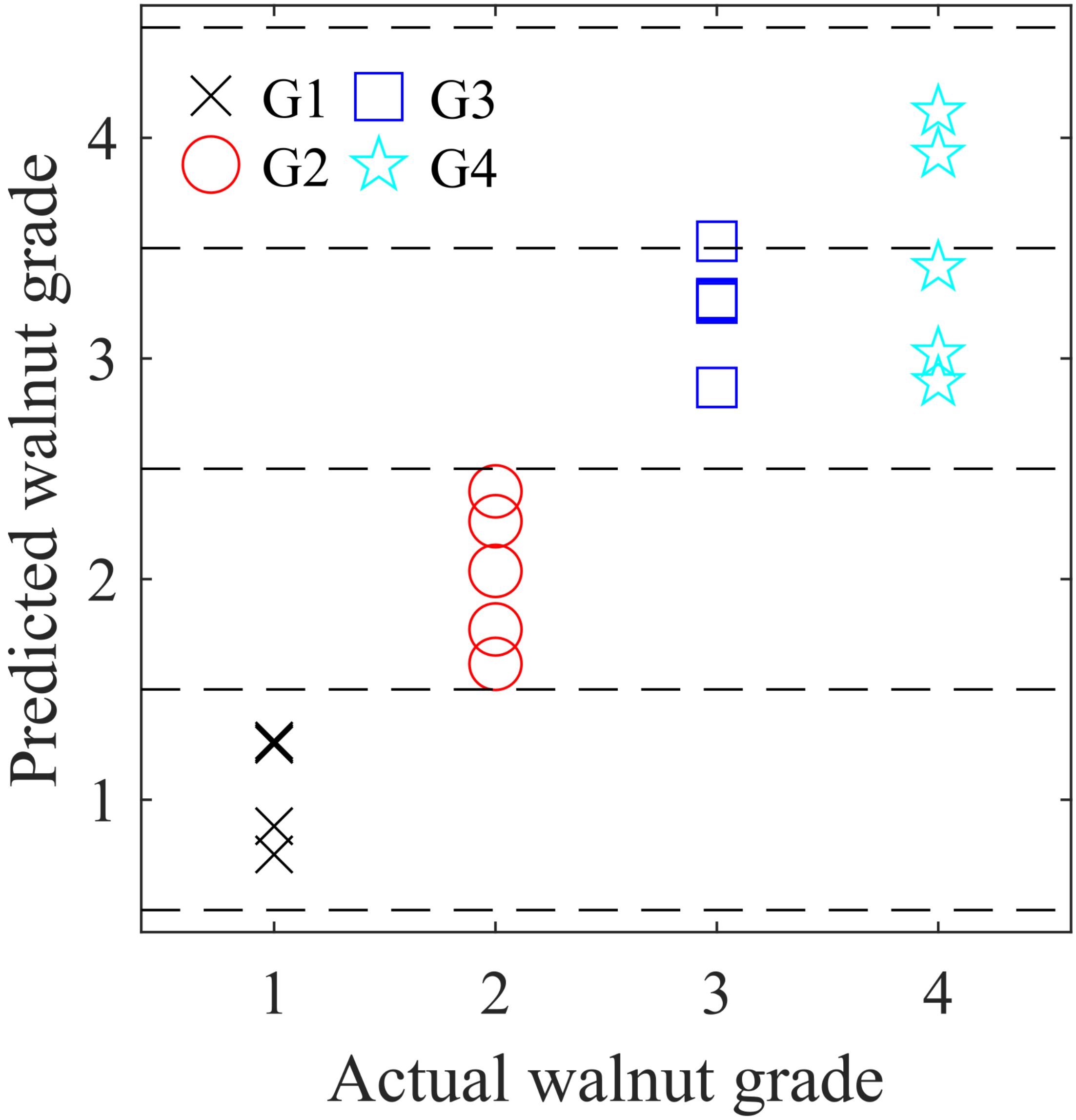

Figure 5:

Comparison of actual and PLSR model predicted grade of walnuts. Two successive dashed lines set the window for a specific grade. The predicted value of 1 ± 0.5, 2 ± 0.5, 3 ± 0.5, and 4 ± 0.5 are considered to be grades 1, 2, 3, and 4, respectively.

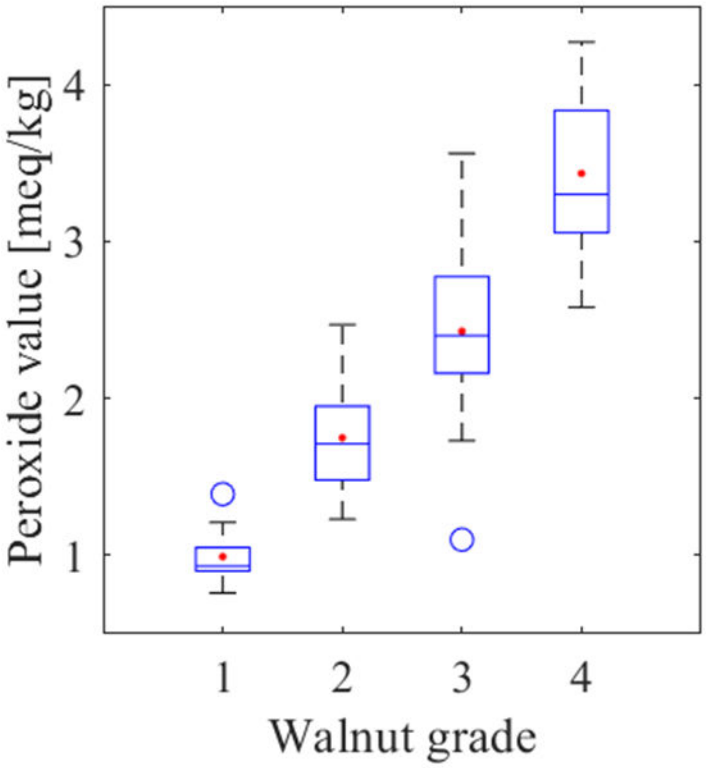

Figure 6:

Variation of Peroxide value as the walnut rancidity changes.

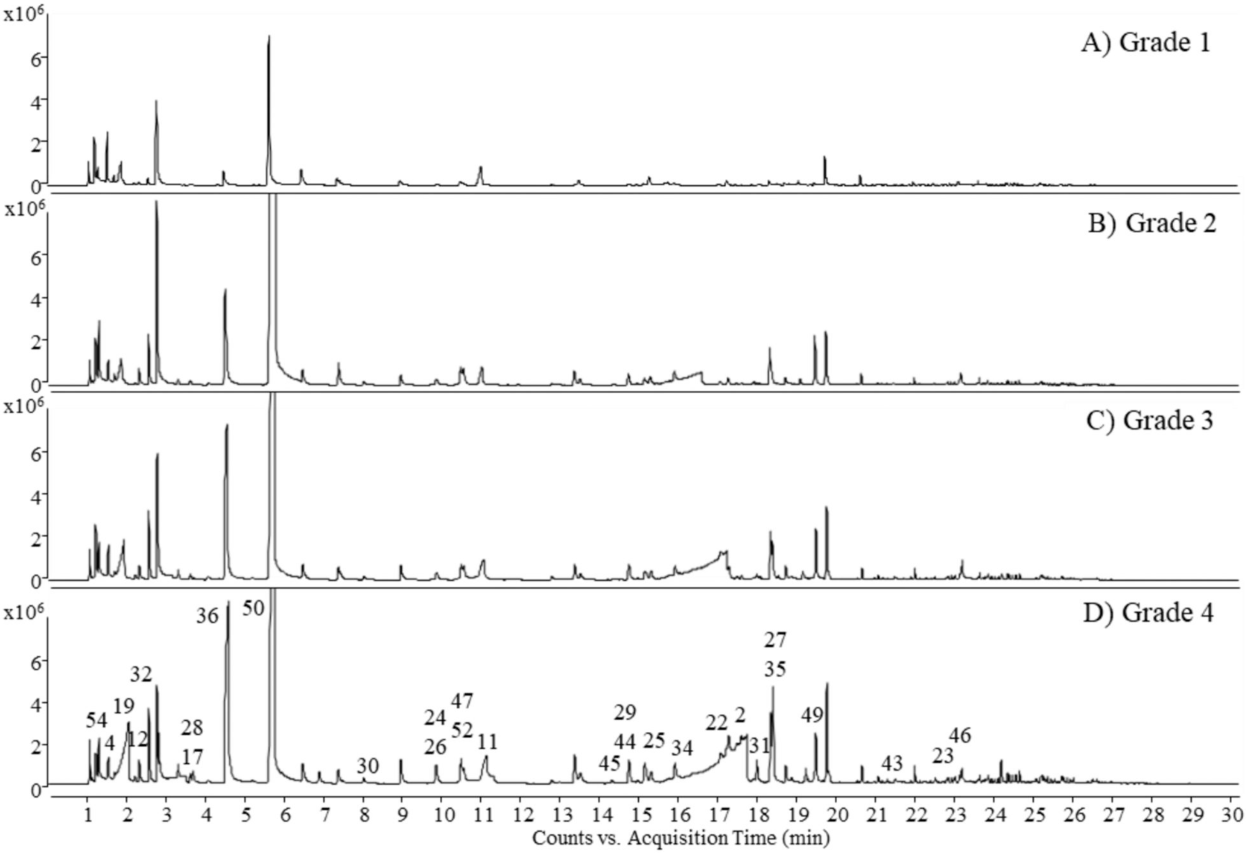

Figure 7:

GC-MS chromatogram of different grades of walnuts. The chemical identity of numbered chemicals can be found in Figure 8.

#	Compound	Formula	Score	Grades			
				G1	G2	G3	G4
1	Methyl 1-methylcyclopropyl ketone	C6H10O	70.6				
2	Hexanoic acid	C6H12O2	77.2				
3	Toluene	C7H8	84.7				
4	n-Hexane	C6H14	83.3				
5	3-Ethyl-2,2-dimethyl-pentane	C9H20	84.3				
6	2,2,4-Trimethyl-hexane	C9H20	72.2				
7	3-Methyl-nonane*	C10H22	91.6				
8	2-Methyl-nonane	C10H22	77.9				
9	2,2,4,6,6-Pentamethyl-heptane	C12H26	65.3				
10	2-Phenylethyl ester acetic acid	C10H12O2	74.9				
11	Methoxy-phenyl-oxime-	C8H9NO2	81.4				
12	3-Methyl-butanal	C5H10O	84.1				
13	2,6,10,14-Tetramethyl-heptadecane	C21H44	78.8				
14	2,5,9-Trimethyl-decane	C13H28	77.0				
15	3,4-Dimethyl-heptane	C9H20	77.3				
16	alpha-Butyl-cyclopropanemethanol	C8H16O	73.9				
17	2-Methyl-2-butenal*	C5H8O	91.7				
18	Nonyl prop-1-en-2-yl ester carbonic acid	C13H24O3	68.8				
19	Acetic acid*	C2H4O2	93.2				
20	2-Methyl-2-pentenal*	C6H10O	93.6				
21	5-Propyl-decane	C13H28	87.9				
22	2-Ethyl-1-hexanol	C8H18O	78.0				
23	Cyclodecanone	C10H18O	85.3				
24	2-Heptanone	C7H14O	84.9				
25	2-Pentyl-furan*	C9H14O	96.5				
26	5-Methyl-2-hexanone	C7H14O	80.5				
27	(E)-2-octenal*	C8H14O	94.9				
28	3-Methyl-1,2-cyclopentanediol	C6H12O2	78.3				
29	1-Octen-3-ol*	C8H16O	96.8				
30	(E)-2-hexenal	C6H10O	88.9				
31	5-Ethyldihydro-2(3H)-furanone*	C6H10O2	96.0				
32	1-Penten-3-ol*	C5H10O	90.8				
33	3,5-Octadien-2-ol	C8H14O	80.5				
34	Octanal	C8H16O	82.3				
35	Anhydride 2-methyl-pentanoic acid*	C12H22O3	90.5				
36	1-Pentanol*	C5H12O	93.7				
37	cis-1,2-Diethyl-	C8H16	85.3				
38	2,6,10-Trimethyl-	C15H32	77.9				
39	1-(2,2-Dimethylcyclopropyl)-2-	C8H12O	67.4				
40	Cyclobutyl heptyl ester oxalic acid	C13H22O4	68.3				
41	(E)-2-pentenal*	C5H8O	93.1				
42	Cyclohexanol	C6H12O	74.6				
43	Octanoic acid	C8H16O2	81.9				
44	2-Hexenal*	C6H10O	97.2				
45	1-Heptanol*	C7H16O	90.3				
46	3-Nonen-2-one	C9H16O	81.8				
47	(E)-2-heptenal*	C7H12O	97.1				
48	Hexyl-oxirane	C8H16O	83.7				
49	2-Butyltetrahydro-furan	C8H16O	85.1				
50	Hexanal	C6H12O	88.6				
51	1-Propenyl-cyclopropane	C6H10	72.3				
52	Pentyl-oxirane	C7H14O	87.3				
53	4-Ethylcyclohexanol	C8H16O	68.5				
54	Propanal	C3H6O	78.4				
55	Trimethyl-oxirane	C5H10O	81.6				
56	Heptanoic acid	C7H14O2	78.8				
57	Butanal	C4H8O	74.5				
58	2-Ethyl-cyclobutanol	C6H12O	82.4				
59	2-Chloro-2-methyl-	C4H9Cl	66.5				

Figure 8:

List of compounds detected in walnuts using HS-SPME-GC-MS. Compound names are listed with their corresponding formula, and score value (in %, based on spectra matching NIST library). Heatmap show changes of the individual compounds intensities for the fore grades defined by storage times. Intensity values are relative to the maximum intensity of the same volatile detected across all samples. Dark blue color represents lower values (0) and dark orange represents higher volatile intensities (1). Non-detected compounds are represented in grey color. Compounds with mass spectral match scores higher than 90% are highlighted with (*).

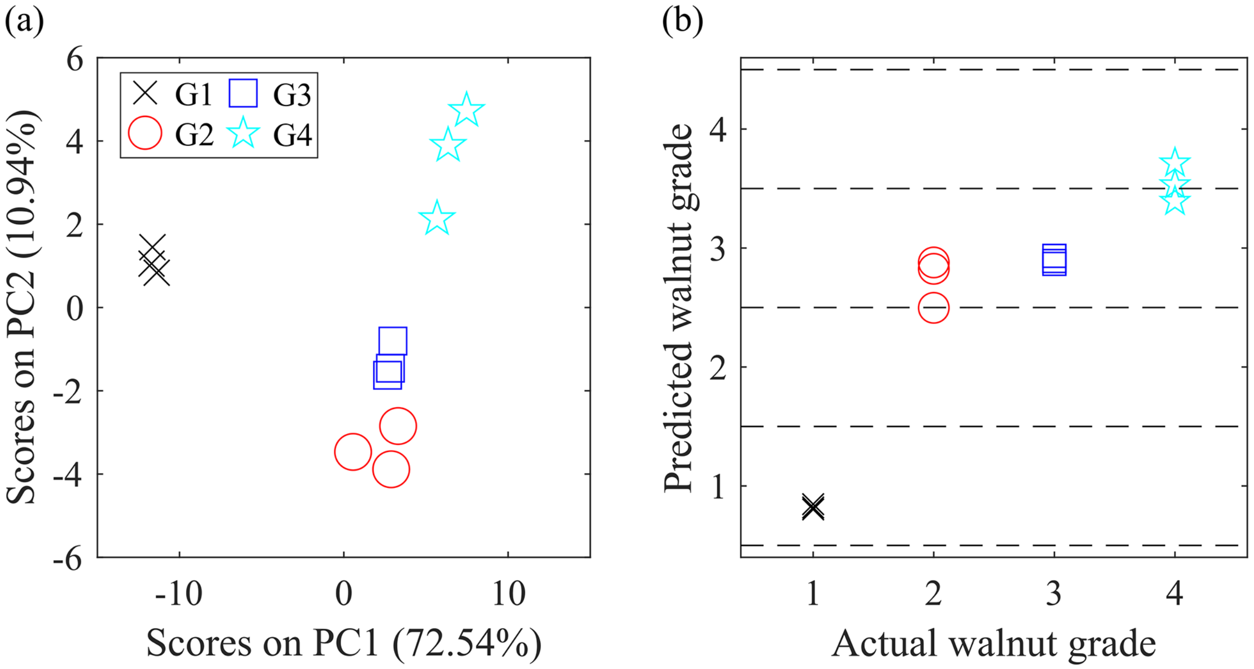

Figure 9:

(a) Scores plot of the first and second principal components resulting from the principal components analysis of the GC-MS data of different grades of walnuts. G1, G2, G3, and G4 represent grade 1, grade 2, grade 3, and grade 4, respectively. (b) Comparison of measured and PLSR model predicted grade of walnuts.

Table 1:

Description about different grades of walnuts. All grades have the same in-shell storage timing and different rancidity appears as a result of different as-kernel storage timing.

Walnut grade	Storage in shell (months)	Storage as kernels (months)	Total storage (months)	Available walnuts (grams)	Experimental Sample size (grams)
Grade 1	5	6	11	~180	30
Grade 2	5	9	15	~180	30
Grade 3	5	12	17	~180	30
Grade 4	5	15	20	~180	30