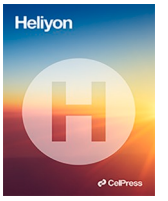
Heliyon 10 (2024) e35687

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Characterization of semi-volatile compounds in 56 Italian ciders using GC×GC-TOF-MS and multivariate analysis

Ciro Orecchio [a], Andrea Bedini [b], Monica Romagnoli [c], Sebastiano Pantò [d], Eugenio Alladio [a], Marco Pazzi [a,*]

[a] Department of Chemistry, University of Turin, Via P. Giuria 7, 10125, Torino, Italy
[b] Founder and member, Associazione Pommelier e Assaggiatori Sidro, A.P.A.S., Torino, Italy
[c] Department of Chemical, Pharmaceutical and Agricultural Sciences, University of Ferrara, Via L. Borsari 46, 44121, Ferrara, Italy
[d] LECO European Application and Technology Center (EATC), Biotechpark, Building B 5.2 Max-Dohrn-Str. 8-10, 10589, Berlin, Germany

ARTICLE INFO

Keywords:
Italian cider
SVOCs profiling
SPME
Comprehensive two-dimensional gas chromatography
Mass spectrometry
Chemometrics

ABSTRACT

Fifty-six samples of differently produced commercial Italian ciders were analysed for semi-volatile organic compounds (SVOCs) profiling, using comprehensive two-dimensional gas chromatography coupled to mass spectrometry (GC×GC-TOF-MS) technique for the very first time. To properly support the compositional investigation of this emerging beverage, a chemometric approach through Principal Component Analysis (PCA) was employed. This revealed a sample distribution in agreement with results of the sensory tasting panel performed on such ciders, highlighting an excellent correlation between variables and perceived odorants. In particular, the positions of peculiar and anomalous objects in the Principal Components (PCs) space are explicitly evaluated, exploring the associated loadings (i.e., the importance of the identified chemical compounds), paying attention to their biochemical origin along the cider-making process and their impact on the sample olfactory analysis. Besides this, the t-distributed Stochastic Neighbor Embedding (t-SNE) method was shown to be an efficient tool for gathering pear ciders from the other samples (apple ciders), better than PCA. This study stands for the first survey on Italian commercial craft cider, and its results are aimed to be a milestone for its characterization and to start and promote cider culture in this country.

1. Introduction

Cider is a traditional low-alcoholic beverage realized by partial or complete fermentation of juice, with or without the addition of sugar, water, or flavouring [1,2]. It is typically produced from apple (*Malus domestica*, Borkh), a premier temperate fruit grown worldwide. Depending upon the fermentation process, the final product will have, in general, an alcohol content between 1.2 % and 8.5 % alcohol by volume (ABV) and should keep the character of fermented apple juice [3]. Moreover, although there are a few apple cultivars that are sometimes fermented into good quality single-cultivar ciders, most ciders are blended from different cultivars to achieve the desired balance of acidity, sugar, and tannins [4]. For this reason, generally, apple ciders can be characterized by different alcohol, sugar, tannins, and CO_2 content, resulting in dry or sweet, bitter or astringent, sparkling or still ciders [5].

Cider is one of the oldest known beverages with a long and fascinating history. Historians broadly agree that apple trees existed

* Corresponding author. Department of Chemistry, University of Turin, Via Pietro Giuria 7, 10125, Torino, Italy.
E-mail address: marco.pazzi@unito.it (M. Pazzi).

https://doi.org/10.1016/j.heliyon.2024.e35687
Received 27 March 2024; Received in revised form 25 July 2024; Accepted 1 August 2024
Available online 4 August 2024

along the Nile River Delta as early as 1300 BCE, and several written documents citing alcoholic beverages made from apples date back to more ancient times in central-east Asia [6]. A preliminary description of cider-making in the Mediterranean area is recorded in the works of the Roman writer Pliny during the first century (A.D.). Then, its production seems to have moved toward northern countries, so that cider-making was well established in France by the Charlemagne's era (9th century A.D.) and perhaps was introduced in England from Normandy well before Duke William's conquest in 1066 [2,5,7,8]. In the late 1980s′, modern ciders were born: ciders with juice and flavourings began to be produced alongside traditional ciders. Since 1990, cider production and consumption have increased above all in France, the United Kingdom, Spain, and the USA, with the increase in craft breweries [4].

Cider can be classified into different styles, each defined by specific cultural practices and production methods. There is not a unique classification, but it is common practice to group them into two main families: Standard Ciders (such as New World Cider, English Cider/West country, French Cider, Basque Cider, Applewine) and Specialty Ciders (such as New England Cider, Flower and Fruit Flavoured Cider, Herbs and Spice Flavoured Cider, Ice Cider). Moreover, also pear cider has its own categories: New World Perry, Traditional Perry, Specialty Perry [9,10].

Over the last twenty years there has been a progressive improvement in the cider processing techniques employed by the industry, including but not limited to harvesting, crushing, fermentation, bottling and post-treatment methods [8,11,12].

The rising knowledge on biochemical phenomena happening along the cider production process pushed the scientific community to deepen the influence of identified organic molecules on the cider sensory profile [13]. Research into a range of topics concerning apple and cider is and has been conducted by many institutes and universities around the world: with the increase in production and consumption of cider, research has also been promoted and the number of scientific documents published annually raised from 50 to 250, in the last twenty years [14]. As all the steps of cider production, from orchard varieties to bottles stored in shelves, are involved in volatile compounds production and transformation, one of the primary studies being carried out is concerning semi-volatile organic compounds (SVOCs) profiling. This is due to their massive ability to affect the quality and the aroma of the finished product.

Usually, the characterization of the semi-volatile compounds is performed by SPME, but Zhang et al. conducted a study using different fermented beverages, including cider, to investigate different types of extraction techniques [15]. In terms of analysis technique, however, the use of gas chromatography both with MS [16–20] and FID [21–23] detectors is well-established. What is still not widely used in the characterization of SVOCs in cider, but is only present in two works, is the comprehensive two-dimensional gas chromatography coupled with mass spectrometry, GC×GC-MS [15,18]. Combining third dimension which is mass spectrometry with GC×GC system generates one of the most powerful analytical tools for semi-volatile organic compounds in complex matrices [24]. It differs from one-dimensional GC in that the sample undergoes two consecutive chromatographic separations with columns of different polarity. This is possible thanks to the modulator, which captures analytes from the first dimension, focusing and releasing them in the second dimension. Comparing the one-dimensional GC and the comprehensive two-dimensional GC, there are many advantages in the two-dimensional one, in terms of separation power, sensitivity, structure, selectivity and speed of analysis [25].

Furthermore, there is currently a lack of cultural and general scientific information about commercial Italian cider; indeed, only a few scientific papers are dedicated to this beverage and not related to any SVOCs fingerprint identification [26–29].

All these circumstances, in addition to the fact that the sensory profile of cider tends to be extremely different from one country to others, pushed authors to explore this field, starting with profiling SVOCs of 56 Italian ciders produced in 2021, using HS-SPME technique coupled with comprehensive GC×GC-TOF-MS methodology, combined with multivariate analysis.

2. Materials and methods

2.1. Cider samples

Italian cider samples were provided from the cultural association 'Associazione Pommelier e Assaggiatori Sidro'. In Table 1 (Supplementary Material) the sample list is provided, together with some information such as cider type, the region where they were produced and some peculiar notes. All cider samples were analysed in triplicate.

2.2. Chemicals and solvents

The chemicals used were NaCl (ACS reagent, ≥99.0 %), obtained from Merck (Darmstadt, Germany) and acenaphthene-d_{10} (CRM) from Restek (Cernusco sul Naviglio, Milano, Italy).

2.3. Samples preparation and extraction

1 g of NaCl and 3 μL of internal standard (acenaphthene-d_{10}, 2 g/L) were added to 5 mL of each cider sample in a 20 mL vial for SPME analysis. Warning: Always follow safety prescriptions using hazardous chemicals.

The extractions were carried out with an autosampler L-PAL3 (LECO Corporation, Saint Joseph, Michigan, USA) using a Merck Stableflex DVB/CAR/PDMS fibre (50/30 μm) at 60 °C for 5 min, after 15 min of incubation at the same temperature, under stirring. DVB/CAR/PDMS coating was chosen as it combines the absorption properties of the liquid polymer with the adsorption properties of porous particles, having macro (>500 Å), meso (20–500 Å) and microporous (2–20 Å); it also shows bipolar properties suitable for a molecular weight range from 40 to 300 Da [16].

2.4. Instrumental parameters

The analyses were conducted on a Pegasus BT 4D GC×GC-TOF-MS equipped with a quad-jet dual-stage cryogenic modulator (LECO Corporation). The chromatographic columns were 30 m × 0.25 mm × 0.25 μm d_f Rxi-5SilMS and 1 m × 0.25 mm × 0.25 μm d_f (0.6 m coiled in the secondary oven) Rxi-17SilMS (both from Restek) as the first and second dimension (^{1}D and ^{2}D), respectively. The analytes were desorbed from the fiber for 2 min at 250 °C in split mode injection (1:50); the carrier gas was helium, used in constant flow mode of 1.4 mL/min. The oven temperature program was 40 °C (held for 2 min), then ramped at 5 °C/min to 230 °C and finally ramped at 25 °C/min to 280 °C (held for 1 min). The temperature offset for the secondary oven and the modulator were set at +5 °C and +20 °C with respect to the main oven temperature, respectively. A 2.5 s of modulation period was used. The transfer line and the ion source temperatures were set at 300 and 250 °C, respectively. A mass range from 35 to 450 *m/z* was collected with an acquisition rate of 200 spectra/s.

Data were collected and analysed using ChromaTOF® BT software version 5.55.41 and ChromaTOF® Tile software version 1.01 (LECO Corporation), respectively. The second software is a tile-based Fisher-ratio (F-ratio) software. Briefly, the chromatograms were divided in regions, tiles precisely, where a signal-to-noise ratio (S/N) threshold was applied, and the F-ratios were calculated at each tile for each *m/z*. The average F-ratio for each tile was calculated, and it was used to rank the values in a hit list where the compounds were identified with the NIST11 library [30,31]. This study used tiles of 7 modulations as ^{1}D size and 61 spectra as ^{2}D size; the S/N threshold was set at 10, and it was considered a 40–250 *m/z* range. The library search was carried out according to the default parameters of tiles and the comparison of retention indices. In Fig. 1 a two-dimensional chromatogram of an analysed cider sample is shown.

The environmental impact of this novel method has been examined with Complex green analytical procedure index [42] together with its applicability e.g., with Blue applicability grade index [43]. Results are shown in Fig. 1S in Supplementary Material.

2.5. Cider tasting

Ciders were analysed by a trained panel of persons, including certified Pommeliers, during a tasting session. Their organoleptic properties were evaluated, including smell and taste intensities and aromatic notes. These observations have been used to support the results of the multivariate analysis; at the same time, SVOCs recognition by GC×GC-TOF-MS technique was of precious importance for understanding the aroma notes identified by the tasting panel. All the participants involved gave informed consent, and they were able to withdraw from the survey at any time without giving a reason. The products tasted were commercially available, thus being allowed for consumption.

2.6. Multivariate analysis

Multivariate statistical approaches play a crucial role, allowing the extrapolation of meaningful information from the data. Statistical analysis and chemometrics were conducted with Python version 3.11.8 [32].

The original dataset of measurements was refined, removing outlier and non-compliant samples, and creating a data matrix made of 155 variables (identified compounds) and 161 objects (cider samples, including triplicates). The criterion applied for this refinement was based on visual inspection of a screening Principal Component Analysis (PCA), to highlight the not-aligned triplicates. Thereafter, concentration values were also checked singularly to assess the final number of objects.

In this dataset, each compound area was converted to the corresponding concentration as μg/L (*ppb*) via normalization with the internal standard area (acenaphthene-d_{10}) and proper dilution factor. This dataset will be hereinafter marked as 'DATASET1'. Subsequently, a reduced dataset was defined, where the most peculiar samples were removed to highlight and explore the distribution of the other samples; this dataset will be named 'DATASET2'.

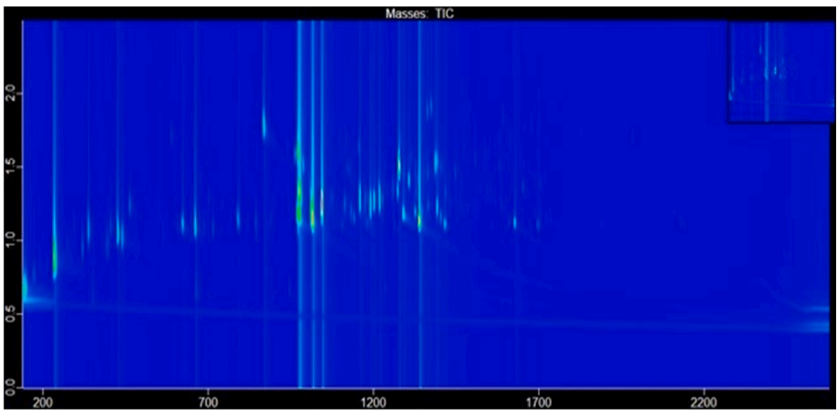


Fig. 1. Example of a two-dimensional chromatogram, related to Sample 56.

The first technique used was PCA, probably the primary unsupervised method of Machine Learning to explore the data. This technique works on p-dimensional variable space used to describe the objects by the X matrix ($n \times p$) to obtain a new, orthonormal vector system, defined by Principal Components, with hierarchization of the unit vectors according to the variance expressed by the data in each of their directions [33,34]. Data were autoscaled before the chemometric treatment.

Secondly, t-Distributed Stochastic Neighbor Embedding (t-SNE), a non-linear, unsupervised, and manifold-based Machine Learning method that performs a dimensionality reduction, was computed [35]. It lends itself particularly well to embedding high-dimensional data in a low-dimensional space (typically 2 or 3 dimensions) while preserving the significant structure of the original data. The algorithm underlying this technique first applies Stochastic Neighbor Embedding (SNE) to the original data, converting the high-dimensional Euclidean distances between data points into conditional probabilities that can be interpreted as similarities. Afterwards, to avoid 'overcrowding', it uses a Student t-distribution (with one degree of freedom) rather than a Gaussian to compute the similarity between two points in the low-dimensional space [36,37] (see S.M. for a more accurate description).

3. Results and discussion

3.1. Most abundant SVOCs profiles in Italian ciders

The GC×GC-TOF-MS analysis carried out a pool of 155 identified compounds that can be clustered in one big group commonly known as "semi-volatile organic compounds", covering several different classes of chemical compounds – alcohols, cyclic hydrocarbons, esters, and terpenes. Some of these components originate from apples, but most are formed during fermentation or derived from added flavourings [11,38].

The concentration of the most abundant SVOCs in Italian ciders is reported in Table 1 and their average molecular weight is around 160 Da, thus correctly aligned with the adsorption range of the fibre employed in this research (see paragraph 2.3). Their concentration spans four orders of magnitude, from few ppb to ten ppm, mainly due to different apples, fermenting and conditioning techniques, and eventual flavourings. Furthermore, the analysis of triplicates pointed out a reliable precision of the estimate for each analyte, with a relative standard deviation average value of 13 % for all the variables.

3.2. Multivariate analysis

PCA was performed on two different datasets ('DATASET1' and 'DATASET2'): in the second one, few samples were removed selectively to highlight the positions of other objects in the PCs space and to explore new correlations with given variables in a more comprehensive way. The positions of the objects in the PCs plane and the related loadings are commented by considering the results of the organoleptic analysis performed on such cider samples.

Table 1
Concentration (μg/L) of main SVOCs in analysed Italian ciders. For chemical structure formula and CAS number see Table 2S in Supplementary Material.

Major SVOCs	Min	Max	Mean	Dev.st
Acetic acid, 2-phenylethyl ester	85.32	4295.26	1003.19	1029.30
Octanoic acid	21.41	3080.68	760.72	686.59
1-Hexanol, 2-ethyl-	40.98	6737.96	355.84	994.25
Hexanoic acid, 2-phenylethyl ester	4.34	7030.71	346.81	1047.23
Ionone	7.46	1564.59	333.67	317.13
Karahanaenone	81.20	1489.26	329.78	259.70
R-Limonene	43.97	1771.02	307.62	322.72
1H-Indene, 2,3-dihydro-1,1,5,6-tetramethyl-	8.80	9645.94	289.64	1312.75
1,3-Octanediol	29.80	1006.90	275.41	237.21
β-Pinene	48.98	3749.23	270.71	565.33
Linalool	31.35	2278.57	248.45	339.37
Butanoic acid, heptyl ester	27.15	987.88	182.89	187.77
Propanoic acid, 2-methyl-, 2-methylbutyl ester	13.74	2798.71	163.62	415.09
2-Undecanol	17.28	1193.33	141.95	185.02
Benzaldehyde	3.77	2736.53	131.47	363.40
α-Terpineol	19.06	1738.19	122.13	277.40
Naphthalene, 1,2,3,4-tetrahydro-1,4,6-trimethyl-	3.06	4346.04	116.49	584.01
1-Octen-1-ol, acetate	14.26	1413.02	114.40	208.98
3-Hexen-1-ol, acetate, (E)-	12.21	700.76	113.44	125.09
5-Hepten-2-one, 6-methyl-	21.64	856.56	106.73	126.10
Theaspirane	2.18	3810.12	104.58	510.24
Butylated Hydroxytoluene	3.08	1502.53	104.10	260.61
Naphthalene, 1,2-dihydro-1,4,6-trimethyl-	2.55	4501.55	100.00	597.11
Geranic acid <methyl-> ester	2.72	2005.09	82.29	317.56
Butanoic acid, 2-methyl-, hexyl ester	20.23	767.01	78.78	108.30
2-Nonanone	4.46	868.30	74.73	160.04
2-Butenoic acid, 2-methyl-, ethyl ester	2.82	2939.13	63.70	350.53

3.2.1. DATASET1

PCA performed on 'DATASET1' is reported in Fig. 2 and Fig. 3. The variance explained for the first three PCs is 34 %, 24 %, and 7 %, respectively. In the PC1 vs PC2 Scores plot, Fig. 2 (a), it can be observed that the presence of a few objects (Samples 56, 55, 54 and 6) deeply affects the distribution of the overall samples, compressing most of them around the centre of the axis.

Sample 56 corresponds to the apple cider having the greatest and most complex aroma profile of the whole dataset, characterized by a strong and unconventional smell of kerosene and tar hints. Moreover, the sum of the identified compound areas for this sample is one order of magnitude greater than others. In the first plot, Fig. 2 (a), it is separated from all the other objects and well described via information from PC2. Looking at the loadings pulling this object in such a direction (Fig. 2 (b)), it can be seen that cyclic hydrocarbon compounds are responsible for this shift. In detail, compounds like 2,3-dihydro-1,1,5,6-tetramethyl-1H-indene, 1,2-dihydro-1,4,6-trimethyl-naphthalene, similar derivatives (hereinafter labelled as TDN), and even their tetrahydro-trimethyl derivatives (labelled TTN) are identified [39]. These C13-norisoprenoids are degradation products derived from the hydrolysis or enzymatic oxidative cleavage of megastigmanes and carotenoids present in fruit, whose presence is often concerned with warm fermentation temperatures,

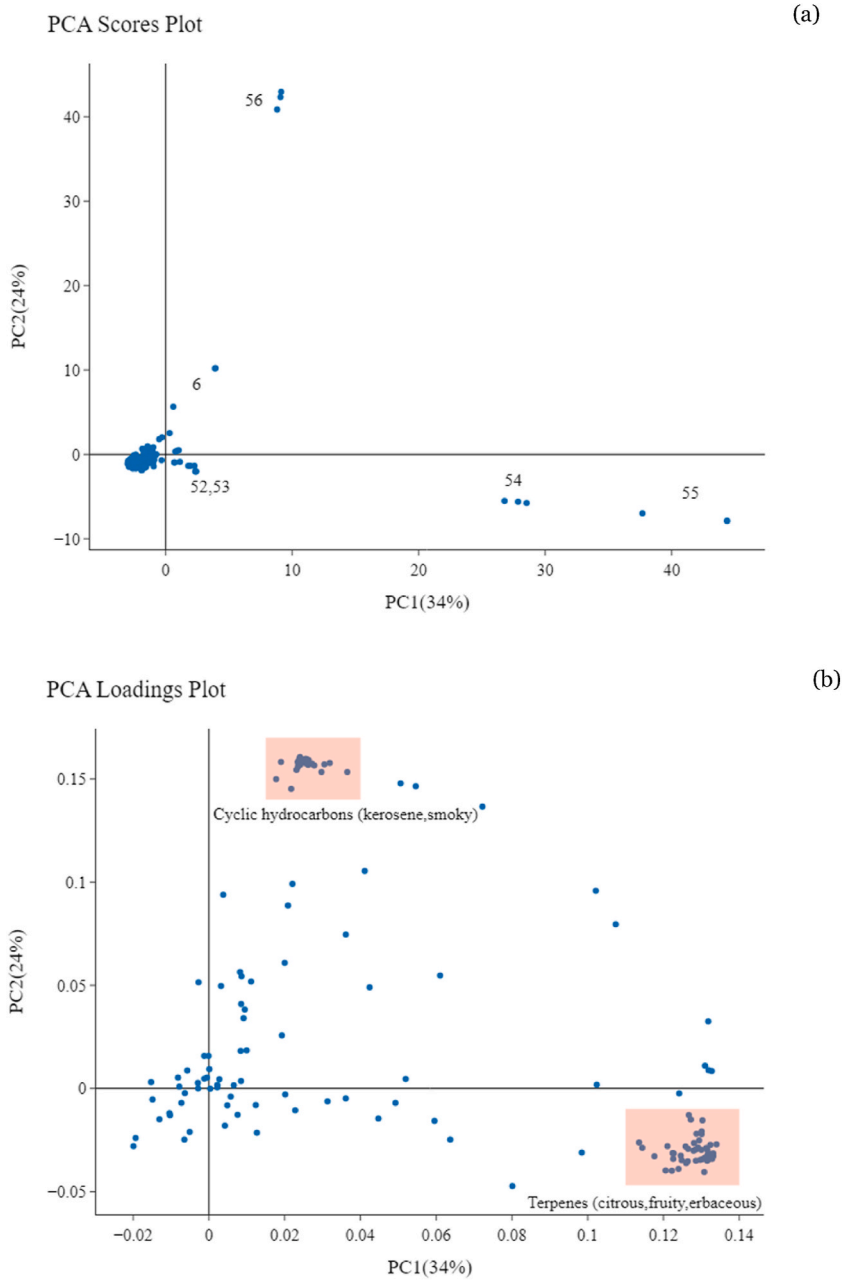


Fig. 2. a) PC1 vs PC2 Scores plot, DATASET1; b) PC1 vs PC2 Loadings plot, DATASET1.

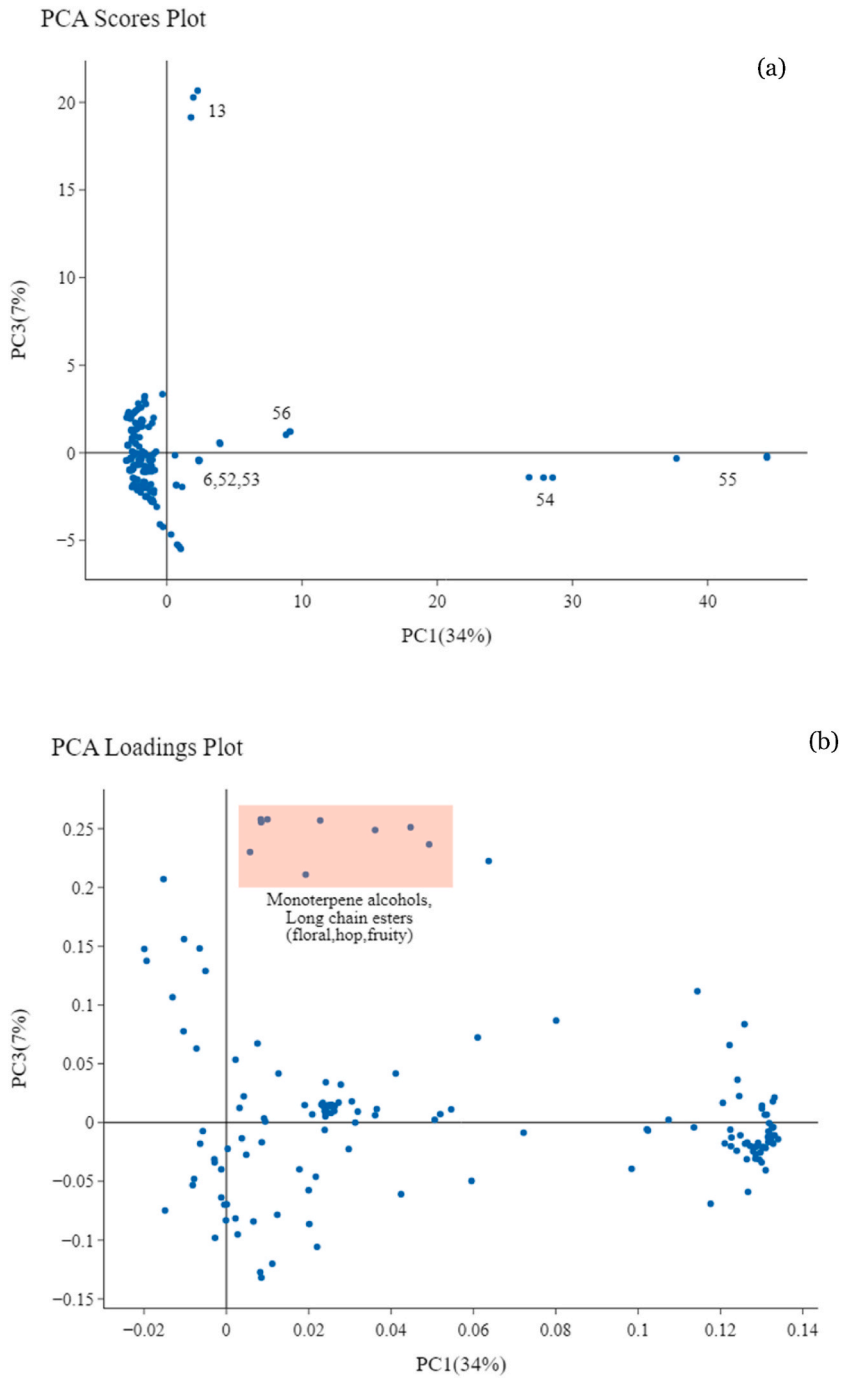


Fig. 3. a) PC1 vs PC3 Scores plot, DATASET1; **b**) PC1 vs PC3 Loadings plot, DATASET1.

air exposure, yeast strains used or unhealthy cork. TDN can provide cider with important aromatics, responsible for smoky, kerosene, and ageing notes. In wine culture, they may be desirable at low concentrations, but may become detrimental when their concentration is very high [40].

Sample 6 is located between Sample 56 and the main group of cider samples, along PC2. The combination of high levels of octanoic acid (rancid), 2-methyl-2-butenoic ethyl ester (aka tiglic acid ethyl ester, earthy-olives), and TDN compounds (hydrocarbons) lead this cider to be judged as defected by the tasting panel. Besides hints of medicine, acetic acid, and over-ripe fruit, a strong note of olives stands out, probably due to the high level of ethyl tiglate, two orders of magnitude higher than all other samples. Moreover, in this case, PCA can highlight an anomalous/defected sample, indeed in the Q-residuals vs. Hotelling's T^2 plot Sample 6 can be interpreted as an

outlier (Fig. 2S; see S.M. for more details).

Samples 54 and 55 are well described by PC1 and correspond to two hopped cider samples, in which hop (*Humulus lupulus* L.) was added via dry-hopping technique to the already fermented cider. The variables related to these objects are indeed typical compounds of essential oils present in hop flowers, such as monoterpenes (α-Pinene), sesquiterpenes (Humulene), and their oxygenated derivatives (Myrceol, Linalool, α-Terpineol, etc). From an organoleptic point of view hopped ciders are really different from pure apple ciders, being characterized by the presence of resinous, herbaceous, citrus, and piney notes both in the aroma and in taste. Samples 52 and 53 are hopped cider samples too, but from the tasting notes it appeared that their intensity was much lower than Samples 54 and 55, according to their position in Scores plot (Fig. 2 (a)), being closer to the rest of the objects, without overlapping them.

PC1 vs PC3 Scores plot, Fig. 3 (a), shows the separation of Sample 13 from the others, along PC3. This sweet cider evidences a distinct flavour, complex and much fruity-floral, with reminiscences of elder, muscat, black currant and lichee. Focusing on the loadings related to this sample (Fig. 3 (b)), it stands out to be rich in long chain and fruity esters (2-phenylethyl acetate, hexenyl and hexyl 2-methyl butanoate), in hop-derived monoterpene alcohols (α-Terpineol and Citronellol), and some compounds typical of essential oils (tetrahydro-4-methyl-2-(2-methyl-1-propenyl)-2H-pyran, 6-methyl-5-hepten-2-one). All these considerations lead us to suppose that this sample is a tailor-made cider, gently hopped, with some fruit syrup or essential oil addition, fermented with selected

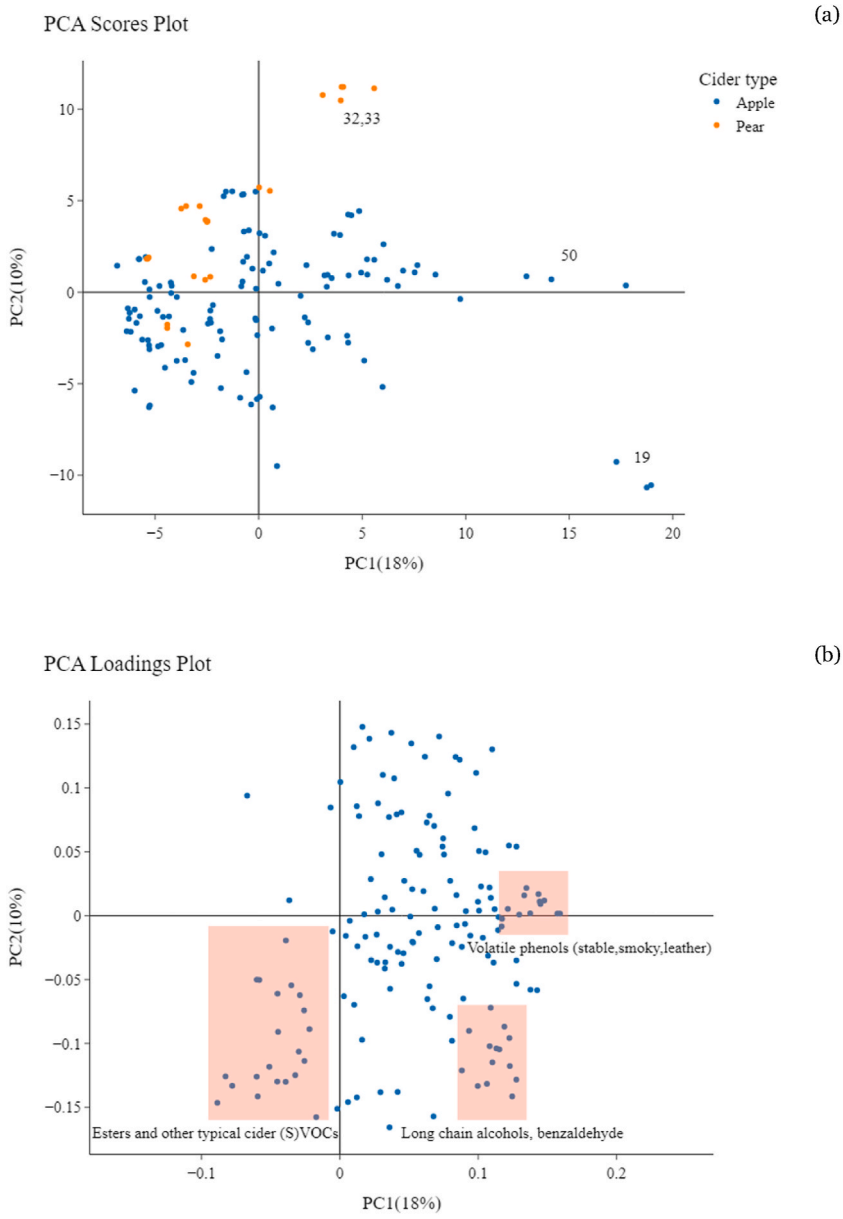


Fig. 4. a) PC1 vs PC2 Scores plot, DATASET2; **b**) PC1 vs PC2 Loadings Plot, DATASET2.

yeast and then pasteurised.

3.2.2. DATASET2

PCA performed on 'DATASET2' is reported in Fig. 4. In this case, because the most peculiar samples were removed, the variance explained by Principal Components is lower than in the previous model: 18 % explained by PC1 and 10 % by PC2. In the Scores plot (Fig. 4 (a)), it can be noticed that there are some peculiar objects (Samples 50, 19, 32 and 33) separated from the cloud of other samples, which here is more defined respect to its aspect on 'DATASET1'.

Sample 50 is detached from others due to its high content in volatile phenols, 4-ethyl-2-methoxyphenol (often referred as 4-ethyl-guaiacol or 4 EG) and 4-(2-propenyl)-phenol. The tasting session outlines a unique aroma rich in leather, stable, smoky, incense notes, which is typical of a strong spontaneous fermentation, and it is in full agreement with the high level of these phenols. These molecules, indeed, are notoriously regarded as markers of spontaneous fermentation operated by the yeast of genus Brettanomyces/Dekkera which can disassemble hydroxycinnamic acids present in the apple juice through a few enzymatic steps [41].

Sample 19 is well separated from the main cluster and is the sample showing the highest content of 2-ethyl hexanol and benzaldehyde. This cider is the only one that came in contact with barrel wood as it results from a fifty-fifty blend of fresh and barrel-aged cider. As the barrel used for this blend is a tonneau also used to age wine, we hypothesize the high content of such two compounds can be derived from a grape must cross-contamination or it can be due to a proper metabolic pathway of the indigenous yeast present in the barrel. This sample showed a distinct complex aroma of rubber, honey, and tar, very likely due to the synergic effect of fermentation inside the barrel and the presence of such spoilage yeast.

Then, other samples start to be detached from the main cluster of objects, such as Samples 32 and 33, which are two pear ciders. To explore more wisely the pear ciders cluster, we preferred to use another technique, t-Distributed Stochastic Neighbor Embedding (t-SNE).

t-SNE, applied on the same dataset (DATASET2), highlights the cluster of pear ciders better than PCA: in the PCA context, a slight separation trend along the bisector of the first and third quadrants can be notice (Fig. 4 (a)). It is interesting to note that Sample 27 does not belong in the cluster with other pear ciders (Fig. 5). This is because in this sample, three apple varieties were fermented together in addition to one pear variety. The t-SNE approach detects this difference and separates efficiently this singular object from the others.

4. Conclusion

The present work is the first to thoroughly analyse and interpret many Italian cider samples. A panel of tasters identified their primary scents; then, their semi-volatile fingerprints using the HS-SPME-GC×GC-TOF-MS technique were detected. Around 130 SVOCs were found in all the analysed ciders, out of a pool of 155 SVOCs that could be categorised into multiple chemical families.

Our primary findings originate from cross-examining the measurements of cider SVOCs and their association with perceived odorants. This led to several results that will interest heterogeneous audiences: Pommeliers and cider expert community from a more technical and chemical perspective, as well as small cider producers and rural farms to master their fermentative processes.

Chemometric methods like PCA were used to emphasise the presence of odd samples found to be anomalous during the tasting session. PCA objects distribution is consistent with the findings of the sensory panel, showing an impressive correlation between sensed odorants and detected variables. Furthermore, t-SNE proved to be more effective in highlighting similarities across different classes of objects (in this case, pear ciders versus apple ciders).

In our scenario, the pairing of two powerful approaches such as bidimensional GC and chemometrics showed to be the best choice to explore and deepen new topics in the food science area, from a compositional and organoleptic point of view.

Moreover, unlike other nations where cider production has a century-old heritage, Italy also produces this historic beverage, making this research paper a promising tool for promoting scientific culture on the subject. Further research, involving new vintages and a wider variety of ciders, will be conducted as these results only offer the fundamentals to understand the typicality of Italian ciders. These studies are essential to verify the correlation between cider molecular odorants and perceived aromatic notes.

Ethics and consent statement

All participants were informed that consent to participate in the study and publish their data would be assumed on completion and submission of the study questionnaire/survey. Moreover, they were able to withdraw from the survey at any time without giving a reason.

Data availability

Data will be made available on request.

CRediT authorship contribution statement

Ciro Orecchio: Writing – review & editing, Writing – original draft, Visualization, Software, Data curation, Conceptualization. **Andrea Bedini:** Writing – review & editing, Writing – original draft, Resources, Investigation, Data curation, Conceptualization. **Monica Romagnoli:** Writing – review & editing, Writing – original draft, Methodology. **Sebastiano Pantò:** Supervision, Resources, Methodology, Investigation. **Eugenio Alladio:** Writing – review & editing, Data curation. **Marco Pazzi:** Writing – original draft,

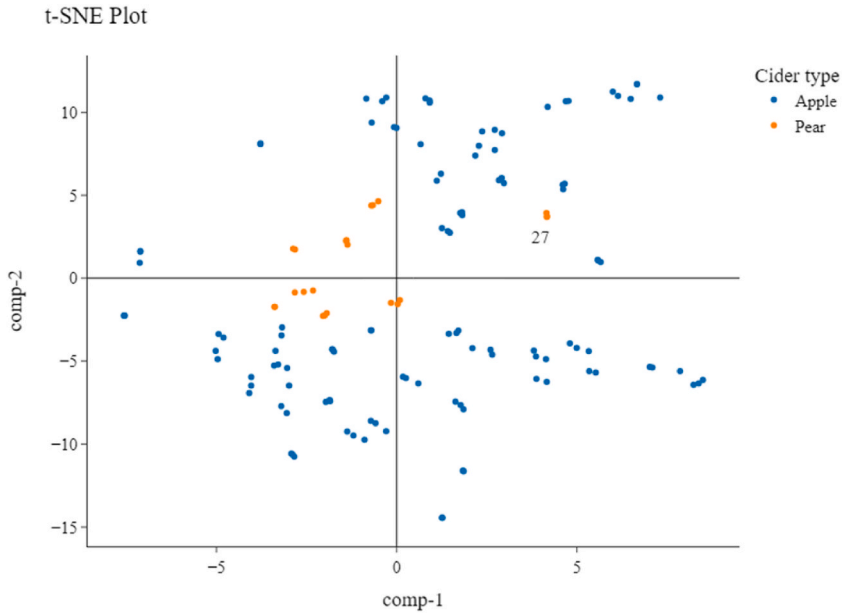


Fig. 5. t-SNE plot, DATASET2.

Resources, Methodology, Investigation, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Ciro Orecchio reports financial support was provided by Project CH4.0. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was carried out within the Agritech National Research Center and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D. D. 1032 June 17, 2022, CN00000022). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

Authors acknowledge support from the Project CH4.0 under the MUR program "Dipartimenti di Eccellenza 2023–2027" (CUP: D13C22003520001).

Authors are grateful to LECO Italy (Dr. Giovanni Bonanomi and Dr. Noemi Cazzaniga) for their precious support.

Special thanks to the cultural association 'Associazione Pommelier e Assaggiatori di Sidro' (https://www.assopommelieresidro.com/) for the chance of sampling for the very first time such a huge number of Italian ciders and collecting precious tasting notes on them. This activity took place within the project of writing the new Slow Food® guide 'Guida alle Birre d'Italia'.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e35687.

References

[1] M. McKay, A.J. Buglass, C. Gook Lee, Cider and Perry, in: Handbook of Alcoholic Beverages: Technical, Analytical and Nutritional Aspects, John Wiley and Sons, 2010, pp. 231–265, https://doi.org/10.1002/9780470976524.ch11.

[2] V.K. Joshi, B.L. Attri, P.S. Panesar, G.S. Abrol, S. Sharma, A.D. Thakur, S. Selli, H. Kelebek, L.V. Reddy, Specific features of table wine production technology, in: Science and Technology of Fruit Wine Production, Elsevier Inc., 2017, pp. 295–461, https://doi.org/10.1016/B978-0-12-800850-8.00007-7.

[3] AICV – European Cider and Fruit Wine Association, European Cider Trends 2022, Brussels, 2022.

[4] I.A. Merwin, S. Valois, O.I. Padilla-Zakour, Cider apples and cider-making techniques in europe and north America, in: Hortic Rev (Am Soc Hortic Sci), wiley, 2008, pp. 365–415, https://doi.org/10.1002/9780470380147.ch6.

[5] V.K. Joshi, S. Sharma, V. Kumar, Cider: the production technology, in: Winemaking, first ed., 2021.
[6] F.J. Cousin, R. Le Guellec, M. Schlusselhuber, M. Dalmasso, J.M. Laplace, M. Cretenet, Microorganisms in fermented apple beverages: current knowledge and future directions, Microorganisms 5 (2017), https://doi.org/10.3390/microorganisms5030039.
[7] F.A. Roach, Cultivated Fruits of Britain: Their Origins and History, Blackwell Pub, Oxford, 1985.
[8] P.C. Calugar, T.E. Coldea, L.C. Salanță, C.R. Pop, A. Pasqualone, C. Burja-Udrea, H. Zhao, E. Mudura, An overview of the factors influencing apple cider sensory and microbial quality from raw materials to emerging processing technologies, Processes 9 (2021), https://doi.org/10.3390/pr9030502.
[9] Gabe Cook, Ciderology, Spruce, 2018.
[10] https://legacy.bjcp.org/docs/2015_Guidelines_Cider.pdf (accessed December 21, 2023).
[11] F.W. Beech, Cider making and cider research: a review, J. Inst. Brew. 78 (1972) 477–491, https://doi.org/10.1002/j.2050-0416.1972.tb03485.x.
[12] A.G.H. Lea, J.-F. Drilleau, Cidermaking, in: A.G.H. Lea, J.R. Piggott (Eds.), Fermented Beverage Production, Springer US, Boston, MA, 2003, pp. 59–87.
[13] A.J. Buglass, Handbook of Alcoholic Beverages, Wiley, 2016.
[14] Scopus , (n.d.). https://www.scopus.com/search/form.uri?display=basic#basic (accessed December 21, 2023).
[15] P. Zhang, S. Carlin, C. Lotti, F. Mattivi, U. Vrhovsek, On sample preparation methods for fermented beverage VOCs profiling by GCxGC-TOFMS, Metabolomics 16 (2020), https://doi.org/10.1007/s11306-020-01718-7.
[16] R. Perestrelo, C.L. Silva, P. Silva, S. Medina, R. Pereira, J.S. Câmara, Untargeted fingerprinting of cider volatiles from different geographical regions by HS-SPME/GC-MS, Microchem. J. 148 (2019) 643–651, https://doi.org/10.1016/j.microc.2019.05.028.
[17] H. Guichard, P. Poupard, L. Legoahec, M. Millet, R. Bauduin, J.M. Le Quéré, Brettanomyces anomalus, a double drawback for cider aroma, LWT 102 (2019) 214–222, https://doi.org/10.1016/j.lwt.2018.12.033.
[18] A. Villire, G. Arvisenet, L. Lethuaut, C. Prost, T. Sérot, Selection of a representative extraction method for the analysis of odourant volatile composition of French cider by GC-MS-O and GC × GC-TOF-MS, Food Chem. 131 (2012) 1561–1568, https://doi.org/10.1016/j.foodchem.2011.10.008.
[19] M.T. Bingman, J.L. Hinkley, C.P. Bradley, C.A. Cole, Aroma profiles of dry-hopped ciders produced with citra, galaxy, and mosaic hops, Appl. Sci. 12 (2022), https://doi.org/10.3390/app12010310.
[20] A.M. Jagatić Korenika, D. Preiner, I. Tomaz, M. Skendrović Babojelić, A. Jeromel, Aroma profile of monovarietal pét-nat ciders: the role of Croatian traditional apple varieties, Horticulturae 8 (2022), https://doi.org/10.3390/horticulturae8080689.
[21] J.M. Le Quéré, F. Husson, C.M.G.C. Renard, J. Primault, French cider characterization by sensory, technological and chemical evaluations, LWT 39 (2006) 1033–1044, https://doi.org/10.1016/j.lwt.2006.02.018.
[22] H. Tangüler, H. Erten, The influence of two yeast strains on fermentation and flavour composition of cider, Flavour Fragr J 37 (2022) 144–153, https://doi.org/10.1002/ffj.3693.
[23] A. Picinelli, B. Suárez, J. Moreno, R. Rodriguez, L.M. Caso-García, J.J. Mangas, Chemical characterization of Asturian cider, J. Agric. Food Chem. 48 (2000) 3997–4002, https://doi.org/10.1021/jf991284d.
[24] L. Mondello, P.Q. Tranchida, P. Dugo, G. Dugo, Comprehensive two-dimensional gas chromatography-mass spectrometry: a review, Mass Spectrom. Rev. 27 (2008) 101–124, https://doi.org/10.1002/mas.20158.
[25] D. Zanella, J.F. Focant, F.A. Franchina, 30th Anniversary of comprehensive two-dimensional gas chromatography: latest advances, Analytical Science Advances 2 (2021) 213–224, https://doi.org/10.1002/ansa.202000142.
[26] F. Mainente, S. Vincenzi, C. Rizzi, G. Pasini, Polyphenols profile and antioxidant activity characterization of commercial north Italian ciders in relation to their geographical area of production and cidermaking procedures, Beverages 9 (2023), https://doi.org/10.3390/beverages9020054.
[27] G. Lomolino, M. Marangon, S. Vincenzi, A. De Iseppi, Sparkling cider paired with Italian cheese: sensory analysis and consumer assessment, Beverages 8 (2022), https://doi.org/10.3390/beverages8040082.
[28] F. Mainente, B. Simonato, G. Pasini, C. Franchin, G. Arrigoni, C. Rizzi, Hen egg white lysozyme is a hidden allergen in Italian commercial ciders, Food Addit. Contam. 34 (2017) 145–151.
[29] F. Mainente, C. Rizzi, G. Zoccatelli, R. Chignola, B. Simonato, G. Pasini, Setup of a procedure for cider proteins recovery and quantification, European Food Research and Technology 242 (2016) 1803–1811, https://doi.org/10.1007/s00217-016-2724-y.
[30] L.C. Marney, W. Christopher Siegler, B.A. Parsons, J.C. Hoggard, B.W. Wright, R.E. Synovec, Tile-based Fisher-ratio software for improved feature selection analysis of comprehensive two-dimensional gas chromatography-time-of-flight mass spectrometry data, Talanta 115 (2013) 887–895, https://doi.org/10.1016/j.talanta.2013.06.038.
[31] B.A. Parsons, L.C. Marney, W.C. Siegler, J.C. Hoggard, B.W. Wright, R.E. Synovec, Tile-based Fisher ratio analysis of comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry (GC × GC-TOFMS) data using a null distribution approach, Anal. Chem. 87 (2015) 3812–3819, https://doi.org/10.1021/ac504472s.
[32] G. Van Rossum, Centrum voor Wiskunde en Informatica Python reference manual, 1995.
[33] I.T. Jollife, J. Cadima, Principal component analysis: a review and recent developments, Phil. Trans. Math. Phys. Eng. Sci. 374 (2016), https://doi.org/10.1098/rsta.2015.0202.
[34] R. Bro, A.K. Smilde, Principal component analysis, Anal. Methods 6 (2014) 2812–2831, https://doi.org/10.1039/c3ay41907j.
[35] F. Anowar, S. Sadaoui, B. Selim, Conceptual and empirical comparison of dimensionality reduction algorithms (PCA, KPCA, LDA, MDS, SVD, LLE, ISOMAP, LE, ICA, t-SNE), Comput Sci Rev 40 (2021), https://doi.org/10.1016/j.cosrev.2021.100378.
[36] L. Van Der Maaten, G. Hinton, Visualizing Data Using T-SNE, 2008.
[37] B. Melit Devassy, S. George, Dimensionality reduction and visualisation of hyperspectral ink data using t-SNE, Forensic Sci. Int. 311 (2020), https://doi.org/10.1016/j.forsciint.2020.110194.
[38] J. Nešpor, M. Karabín, K. Štulíková, P. Dostálek, An HS-SPME-GC-MS method for profiling volatile compounds as related to technology used in cider production, Molecules 24 (2019), https://doi.org/10.3390/molecules24112117.
[39] G. Versinii, A. Rappz, J. Marais, R. Mattivii, M. Spraul, A New 1,1,6-Trimethyl-L,2-Dihydronaphthalene (TDN) Precursor Isolated from Riesling Grape Products: Partial Structure Elucidation and Possible Reaction Mechanism, 1996.
[40] R.S. Jackson, Wine Science, Academic Press, 2008.
[41] N. Buron, M. Coton, P. Legendre, J.Ô. Ledauphin, V. Kientz-Bouchart, H. Guichard, D. Barillier, E. Coton, Implications of Lactobacillus collinoides and Brettanomyces/Dekkera anomala in phenolic off-flavour defects of ciders, Int. J. Food Microbiol. 153 (2012) 159–165, https://doi.org/10.1016/j.ijfoodmicro.2011.11.002.
[42] J. Płotka-Wasylka, W. Wojnowski, Complementary green analytical procedure index (ComplexGAPI) and software, Green Chem. 23 (2021) 8657–8665, https://doi.org/10.1039/D1GC02318G.
[43] N. Manousi, W. Wojnowski, J. Płotka-Wasylkacd, V. Samanidou, Blue applicability grade index (BAGI) and software: a new tool for the evaluation of method practicality, Green Chem. 25 (2023) 7598–7604, https://doi.org/10.1039/D3GC02347H.