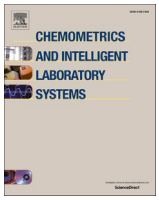
Chemometrics and Intelligent Laboratory Systems 241 (2023) 104940

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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



Optimization of an analytical method based on SPME-Arrow and chemometrics for the characterization of the aroma profile of commercial bread

Samuele Pellacani [a], Caterina Durante [a,*], Silvia Celli [b], Manuel Mariani [b], Andrea Marchetti [a], Marina Cocchi [a], Lorenzo Strani [a]

[a] Department of Chemical and Geological Sciences, University of Modena e Reggio Emilia, Via Campi 103, Modena, Italy
[b] Barilla G. e R. Fratelli, Via Mantova 166, 43122, Parma, Italy

ARTICLE INFO

Keywords:
Bread
Solid phase microextraction arrow
Volatile organic compounds
PARADISe
GC-MS

ABSTRACT

A SPME-Arrow GC-MS approach, coupled with chemometrics, was used to thoroughly investigate the impact of different types of yeast (sourdough, bear's yeast and a mixture of both) and their respective leaving time (one, three and five hours) on VOCs of commercial bread samples. This aspect is of paramount importance for the baking industry to adjust recipe modifications and production parameters, as well as to meet consumer needs in formulating new products.

A deep learning approach, PARADISe (PARAFAC2-based deconvolution and identification system), was used to analyse the obtained chromatograms in an untargeted manner. In particular, PARADISe, was able to perform a fast deconvolution of the chromatographic peaks directly from raw chromatographic data to allow a putatively identification of 66 volatile organic compounds, including alcohols, esters, carboxylic acids, ketones, aldehydes. Finally, Principal Component Analysis, applied on the areas of the resolved compounds, showed that bread samples differentiate according to their recipe and highlighted the most relevant volatile compounds responsible for the observed differences.

1. Introduction

Aroma is one of the most salient characteristics of bread linked to the type of yeast and leavening process. In an industrial context, high-quality production requires that final products maintain the same quality and flavour expected by consumers. This issue is particularly important when production is handled in different countries and plants, using various raw materials. However, the assessment of bread quality is a complex process, and the industry usually relies on evaluating bulk parameters such as colour, taste, smell (through sensory analysis), volume and texture. Aroma is among the most important parameters influencing consumer choice and loyalty, and its assessment, in the case of commercial bread, is carried out through a panel test. Therefore, a deep chemical characterization of Volatile Organic Compounds (VOCs) in bread is of utmost importance for evaluating its quality and supporting sensory evaluation. To characterize the food flavour pattern, numerous analytical methods have been developed, including essential oil extraction with solvents and the collection of the released volatile molecules using several analytical tools [1,2].

In the field of solid phase microextraction, a new device, namely SPME-Arrow, has recently been proposed as an extraction technology for the analysis of volatiles in food materials [3]. This device is based on a novel SPME geometry that can efficiently "hunt" target molecules in complex matrices. The Arrow-SPME configuration overcomes limitations of conventional devices (e.g. mechanical robustness, fused silica's physical durability and small extraction phase volume). Previous studies have demonstrated its potential for characterizing VOCs in various food products like fish [4], vinegar [5], milk [6], grape skins [7] and distillates [8].

Furthermore, several researchers have focused on developing analytical techniques to characterize bread's *volatilome* [9–16] as well. This involves optimizing VOCs sampling/extraction conditions, acquiring the respective chromatograms, identifying peaks and quantifying using appropriate standards.

The aroma profile of bread contains over 300 analytes from different chemical classes (carboxylic acids, aldehydes, ketones, alcohols, esters,

* Corresponding author. University of Modena and Reggio Emilia, Via Campi 103, Modena, Italy.
E-mail address: caterina.durante@unimore.it (C. Durante).

https://doi.org/10.1016/j.chemolab.2023.104940
Received 11 May 2023; Received in revised form 21 August 2023; Accepted 21 August 2023
Available online 25 August 2023

etc.) [16]. Their presence and intensity result from the used raw materials and the bread production process, which involves three main steps: ingredient mixing and dough preparation, fermentation, and baking. Among the raw materials used in bread production, the type of yeast employed during dough fermentation, such as yeast beer and/or lactic acid bacteria (LAB), plays a crucial role, along with enzymatic activities and thermal reactions that take place during the bread baking process [17,18]. Yeasts are important in the synthesis of volatile compounds, such as ethanol, CO_2, alcohols, aldehydes, acids, esters and ketones [19]. This plethora of compounds can be obtained from different fermentation processes and pathways, such as Embden-Meyerhof-Parnas pathway, glyoxylate cycle, the tricarboxylic acid cycle (TCA), the Ehrlich and the shikimate pathway [15]. The overall aroma profile also encompasses pyrazines, pyrrolines, hydrocarbons, furans and lactones produced from lipid oxidation and heat-induced Maillard reaction [19].

In this study, the performance of two SPME devices, namely SPME-Arrow and conventional-SPME fiber, was systematically compared in sampling bread aroma. The comparison began with the optimization of the three main parameters, namely exposure time, exposure temperature and sample incubation time, which influence the VOCs sampling performance [8,20,21], using an experimental design technique. Additionally, the difference in terms of fiber blanks, repeatability, and sensitivity between the two devices was assessed. An untargeted approach was used, where the monitored chemical profile, i.e., the TIC (total ion count) chromatogram was used as a fingerprint of the investigated bread samples. In particular, the adsorption/absorption phase of analytes, during sampling, is influenced by competitive phenomena due to several factors, including the concentrations of the analytes in the headspace, their chemical characteristics (i.e. polarity), affinity with the fiber and so on [22,23]. Variations in extraction times and temperatures may influence the kinetic and thermodynamic mechanisms present during the sampling, leading to different chromatographic profiles. Therefore, the use of the TIC signals (untargeted approach) as response to optimize, instead of the total peak area, allowed to also focus on eventual analytes present only in particular sampling conditions. Simultaneously, the aim was to also maximize the intensity of analyte signals as much as possible. While this approach does not involve an a priori assumption of the analytes to be monitored, it requires the use of chemometric data processing techniques to extract useful information from the obtained signals. In the GC-MS technique or in hyphenate techniques in general, each sample is represented by a two-dimensional map (2D signal or landscape), with the chromatographic profile in one dimension and the mass spectrum in the other. The processing of these signals requires high computational capabilities and efficient methods of dimensionality reduction. Therefore, in this work, the multi-way data analysis approach, PARADISe (PARAFAC2-based deconvolution and identification system [24]) was used to make a deconvolution of the chromatographic peaks directly from raw chromatographic data and to simultaneously perform the integration of the areas of the deconvoluted peaks for all samples.

Furthermore, this research also concerned the study of the influence of three types of yeasts, sampled at different leavening times on the final flavour profile of obtained bread samples. A precise understanding of aroma and the factors influencing its variation is crucial for the baking industry. It enables effective adjustments to recipes and production parameters while also meeting consumer demands for developing innovative products. The application of Principal Component Analysis (PCA) on the peak areas (resolved by PARADISe) data set allowed obtaining a preliminary information on the variability of VOCs present in the studied bread samples.

2. Materials and methods

2.1. Samples

For the aims of this study, different industrial soft bread samples, produced in the same pilot plant, were analysed. The same type '0' soft wheat flour was used to produce all the analysed samples. The yeast type and the leavening time used in the production of bread samples are reported in Table 1. The bread samples were stored at −20 °C until volatile organic compounds were determined. The samples were drawn from the middle of bread slices (Fig. A1, *Appendix A, Supplementary Materials*). All the samples were analysed in duplicate, taking two pieces of the middle bread from two separate slices.

The variability of the experimental procedure was evaluated using a saltines sample purchased from a local store as *control sample*. This choice was made due to its greater stability, compared to bread, during storage time, i.e. the time necessary to undertake all experimentation. The control sample was analysed several times, at least once for each experimental session, according to the same analytical procedure used for the investigated samples.

Finally, a sliced soft bread sample, made of wheat flour, was used in the Design of Experiment procedure for the optimization of both investigated fibers. It was purchased from a large-scale retail store, and it was kept in its original package at room temperature. The choice of a commercial sample is due to ensure that there will be enough to undertake all the experiments required by the design, so to keep constant the sample and focus on the effect of the controlled factors.

Before the analysis, all samples were minced by hand with a cutter to obtain as most homogeneous pieces as possible and a ground sample of around 1.0 g was transferred into a 10 mL glass vial with an aluminium top-closure and silicone/PTFE septum (Chromacol).

2.2. Instrumental analysis

As conventional-SPME a Divinylbenzene/Carboxen/Polydimethylsiloxane (DVB/CAR/PDMS) (50/30 μm × 10 mm) fiber with a diameter of 80 μm was chosen. SPME-Arrow fiber was also a DVB/CAR/PDMS (120 μm × 20 mm) but with a diameter of 1.1 mm (Restek Corporation, Bellafonte, USA). Both the fibers were attached to an SPME fiber holder (Supelco) for extraction procedure. Prior to the experimental analysis, both the fibers were preconditioned in the injector port of the GC system, according to temperature/time conditions recommended by the manufacturers.

The VOCs extraction procedure of conventional-SPME and Arrow-SPME fibers was optimised as described in *Section 2.3*. The final best conditions (as described in *Section 3*) resulted to be the same for both the fibers, namely: the weighted sample (around 1 g) was incubated at 50 °C for 10 min and then the SPME fiber was exposed for 30 min.

The fiber was manually transferred to the split/splitless injector of the GC (6890 gas chromatograph, Agilent Technologies, Santa Clara, CA, USA). The desorption step was performed in splitless mode (3 min) and by setting the injector temperature at 260 °C. After 3 min from the injection, a split mode was activated with a split flow of 25 mL/min. Desorption was done for 5 and 13 min for conventional-SPME and Arrow-SPME fibers, respectively. For the latter, a higher desorption time was necessary to overcome the carry-over phenomena, mainly due to the presence of ethanol, which was used in the recipe of the analysed breads to ensure a better storage. A chromatographic Rxi™ -1 ms column (54 m × 0.25 mm ID, 1 μm) by Restek Corporation was used, and helium as carrier gas at a constant flow rate of 1 mL/min. The GC oven temperature was programmed at 40 °C for 1 min, ramped 4 °C/min to 150 °C,

Table 1
Analysed bread samples description in terms of used yeast in the receipt and leavening time.

Sample label	Yeast type	Leavening time
STD_1	Beer yeast	1 h
C_1	Sourdough	5 h
C_2	Beer yeast & sourdough	1 h
C_3	Beer yeast	3 h

then at 8 °C/min to 250 °C held for 19 min. The GC was interfaced with an Agilent 5973 N mass spectrometer. The detection was performed under electron impact (EI) ionisation at 70 eV by operating in the full-scan acquisition mode in the 25–400 *m/z* range. The transfer line was heated to 270 °C.

In each measurement's session, one run of the SPME fiber, without any analyte's extraction, i.e. a *fiber blank*, was performed, in order to monitor the performance of the fiber as well as any carry-over phenomena.

2.3. Experimental design

One of the first steps in the optimization of an analytical method is the choice of the factors to be investigated and their range of variability. On the basis of literature [8,20,21,25], three main factors, namely exposure temperature, T_{exp}, incubation time, t_{inc}, and exposure time, t_{exp}, were selected and investigated for both the fibers. These factors act on different aspects of the VOCs extraction process. In particular, the exposure temperature mainly influences the thermodynamics phenomena of SPME adsorption/absorption mechanisms; the incubation time has to ensure, especially for solid matrix, the equilibrium of analytes between sample and headspace; finally, the exposure time relates to the kinetic of the absorption/adsorption procedure [26]. The three factors were studied at three levels, low (−), high (+) and central (0), following a Face-Centered Central Composite Design (FC-CCD) [27,28]. The FC-CCD design, comprise the experiments of a full Factorial Design at two level to which six-star points are added (one for each face of the cubic domain), and allows estimation of the main and quadratic terms for each factor and their interactions. Thus, seventeen experiments were planned, which consisted of 8 corner experiments, 6 axial experiments and three replicates of the central point. Finally, the different planned chromatographic runs were randomly performed. The values of the two levels (for each factor) and the design matrix (the same for both fibers) are shown in Table 2 and Table 3, respectively.

The obtained TIC profiles collected across the DoE experiments were preprocessed as detailed in *Section 2.4.2*and firstly analysed by PCA [29] in order to highlight the similarity/differences among the experimental runs and how they link to the chromatographic profile. Inspection of the scores and loadings plots, showed that the first principal component (PC1) could hold information about the overall extracted volatile substances. Hence PC1, was used as response (the higher the PC1 scores value the higher the fraction of extracted compounds). Thus, a multilinear regression (MLR) model was built considering as independent variables, X-block, the design matrix augmented by the interaction and quadratic terms, and as a response, **y**, the PC1-scores. The MLR model was then exploited to study the response surface and finding the optimal operative conditions.

2.4. Data analysis

Notation

Scalars are indicated by italic letters; bold lowercase and uppercase letters are used for vectors and matrices, respectively. For sake of clarity, Table 4 provides all the information about the notations, mathematical operations and abbreviations used in this manuscript.

Table 2
Factors and levels used in FC-CCD.

Factor	Abbreviations	Lower level (−1)	Middle level (0)	Higher level (+1)
Exposure temperature (°C)	T_{exp}	30	40	50
Exposure time (min)	t_{exp}	1	15.5	30
Incubation time (min)	t_{inc}	10	15	20

Table 3
Design matrix of the FC-CCD used for the optimization of VOCs extraction.

Sample label[a]	T_{exp}	t_{exp}	t_{inc}	Response[b]/10^6	
				Arrow	Conventional
S1	−1	−1	−1	−5.56	−3.12
S2	1	−1	−1	−3.40	−2.41
S3	−1	1	−1	1.51	1.32
S4	1	1	−1	5.09	3.08
S5	−1	−1	1	−5.91	−2.41
S6	1	−1	1	−3.13	−2.93
S7	−1	1	1	0.72	1.72
S8	1	1	1	3.84	3.12
S9	−1	0	0	−1.13	0.58
S10	1	0	0	1.67	1.29
S11	0	−1	0	−4.11	−2.36
S12	0	1	0	3.09	1.52
S13	0	0	−1	1.27	0.19
S14	0	0	1	2.14	0.77
C1	0	0	0	1.14	0.39
C2	0	0	0	1.29	1.18
C3	0	0	0	1.44	0.89

[a] In the text each experiment referring to Arrow-SPME is labelled by its number preceded by 'S', while the suffix 'T' will be used when referring to run conducted with the conventional-SPME.
[b] Scores value on the first PC calculated on pre-treated TIC profiles.

Table 4
Nomenclature used in the manuscript.

A, B, D, E, H, X	Two-way data and parameter matrices	MLR	Multilinear regression
F, I, J, K	Dimensions of matrices	PCA	Principal component analysis
f, i, j, k	Indexes of the dimensions of the matrices	PARAFAC2	PARAllel FACtor analysis 2
y T_{exp} t_{inc}	Vectors	DoE	Design of experiment
GC-MS	Gas chromatography coupled with mass spectrometry	PARADISe	PARAFAC2-based deconvolution and identification system
SPME	Solid phase micro-extraction	$(\mathbf{B_k})^T$	Transpose of matrix $\mathbf{B_k}$

2.4.1. Data preprocessing

A proper preprocessing of chromatographic signals is of utmost of importance to achieve reliable results; therefore, in this study, the following preprocessing strategy was implemented.

In the first part of the data analysis, i.e. signals from experimental design approaches, the raw chromatograms (made up of 20212 data points each) were imported in Matlab (release 2020a, The Mathworks Inc., Natick, MA, U.S.A.); the first 3630 and the last 6916 data points for both fibers, corresponding to the retention time (R_t) intervals 7 min and 37 min respectively, were cut since these are due to peaks originating from atmospheric absorbed gas ($R_t < 7$ min) and to chemical species released by the fiber itself ($R_t > 37$ min).

The signals corresponding to each fiber were analysed disjointly and each data set was aligned along the R_t dimension, using the *iCoshift* algorithm [30] taking as reference the average TIC signal. The first alignment of the whole signal was followed by an interval-based alignment. The intervals were manually defined (on the TIC signals), in such a way to hold a single or few peaks each, as reported in Fig. A2, *Appendix A, Supplementary materials*. The alignment step is needed to compensate for the retention time shift, which a peak may exhibit among different chromatographic runs, that could introduce variability among samples not imputable to real differences.

Since some of the observed peaks were due to constituents released by the fiber (see *Section 3.1*), a first correction was carried out by subtracting the average of the registered *blanks* chromatograms from each

sample chromatogram. Furthermore, *Automatic Whittaker Filter* (AWF) [31] was applied on the aligned chromatograms in order to remove baseline offset. In particular, it allows baseline subtraction using the Eilers method based on a Whittaker filter [31]. This algorithm is fast and handles more structured baselines correction, however since it could introduce peak shape artefacts the resulting corrected chromatograms were carefully inspected.

Prior to PCA the pretreated chromatograms, were mean centered.

As regards the second part of data analysis, chromatographic signals coming from bread samples, the same initial R_t interval cut was operated while the *iCoshift* alignment was applied inside PARADISe software (being implemented there). In this case the correction for the blank signal was not applied since, as explained in *Section 2.4.2*, the area of the resolved peaks was used in the subsequent PCA instead of the whole chromatogram, thus the peaks due to the fiber were skipped afterwards.

2.4.2. PARAFAC2

PARADISe is based on PARAFAC2 algorithm [32,33] which allows to recover, by unique decomposition, the mass spectra of the pure components present in the investigated sample and the corresponding concentration profile In this way, it is possible to integrate the areas of the chromatographic peaks corresponding to these deconvoluted components for all samples [24]. PARAFAC2 operates similarly to PARAFAC [34] but can cope with departure from trilinearity due e.g., to shift in one of the mode, such as retention time shifting in the case of GC-MS data. The PARAFAC2 decomposition model can be expressed according to Eq. (1):

$$\mathbf{X}_k = \mathbf{A}\cdot\mathbf{D}_k\cdot(\mathbf{B}_k)^T + \mathbf{E}_k;\ k = 1, \ldots, K \qquad \text{(Equ.1)}$$

where $\mathbf{X_k}$ (I x J) represents the VOCs chromatographic signal related to the k_{th} sample. I and J are the m/z and time dimensions, respectively. The element x_{ij} contains the mass over charge intensity measured at mass i at elution time point j, over all I m/z values and J time points. $\mathbf{A}$ (I x F) holds the loadings of the first mode (mass spectra, I), which are common to all K samples, for each of the F factors. These can be the resolved mass spectra of single analyte present in the sample, as well as trends due to baseline, gradient, etc. $\mathbf{D}_k$ (F x F) is a diagonal matrix holding the weights to apply to the matrix $\mathbf{B}_k$ (J x F), which contains the estimated elution profile corresponding to each factor for each k_{th} sample. It is worth noting that weights are equal to the peak areas when PARAFAC2 is applied to GC-MS data. Finally, $\mathbf{E_k}$ is the residual matrix.

Thus, PARAFAC2 can be able to separate mixture data into concentrations, elution profiles and mass spectra of the underlying analytes [32]. Differently from PARAFAC, PARAFAC2 calculates an individual elution profile for each sample, and thus it can handle issues related to retention time shifting, peak overlap, baseline drift and so on. Moreover, an important constraint in PARAFAC2 is that the cross-product matrix $\mathbf{H}$ (Equ. 2) must be constant across all k samples [32] to maintain uniqueness of the solution.

$$\mathbf{H}=(\mathbf{B_k})^T\cdot\mathbf{B_k} \qquad \text{(Equ.2)}$$

The most relevant implication of this cross-product is that only the elution may differ due to shifting while each analyte has a mass spectrum that is consistent across all samples.

2.4.3. PARADISe approach

The VOCs chromatographic signals of all bread samples were arranged as a three-dimensional array having the following three dimensions (Scheme 1): the recorded mass fragment (i.e, mass spectra as first mode), elution profiles (second mode), and bread samples (third mode).

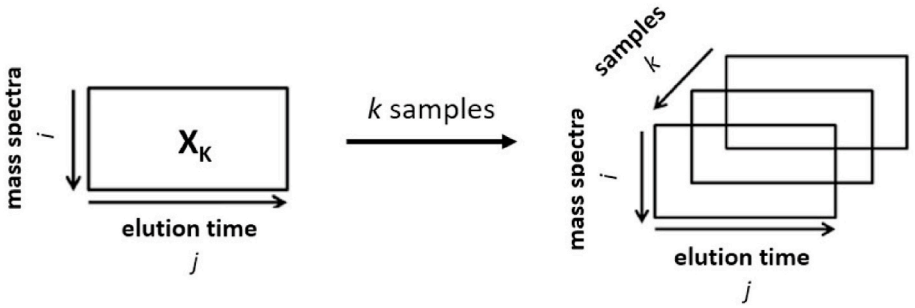
Given to the complexity of data array, the PARADISe method was applied to obtain an efficient and rapid extraction of useful information, i.e., areas of deconvoluted peaks and their putative identification. For computational reasons, applying PARAFAC2 on the whole GC-MS signal is not efficient. Thus, the first step in PARADISe is the definition of retention time intervals to split the signals and applying distinct PARAFAC2 model to each defined interval. In this study, the intervals were manually selected through an interactive TIC graph by using as guideline criterion to have in a single interval a chromatogram portion where, when possible, only one observable peak is present, or at least as few as possible (Fig. A3, *Appendix A, Supplementary Materials)*. Anyhow, PARADISe allows the definition of overlapping intervals and if a peak is resolved in more than one interval it will be possible to select the one where it is best resolved. An example is provided in Fig A4, *Appendix, Supplementary Materials* After defining each interval, separated PARAFAC2 models are automatically calculated for each interval considering models including from one to seven factors. The best dimensionality was chosen according to the explained variance by the model, the entity of residuals, the model's fit (%), the core consistency [35] and the number of peaks identified by each model [36].

The putative identification step is done by PARADISe by checking the similarity between the resolved mass spectrum (corresponding to each of the selected peaks) and the mass spectrum in a database (NIST MS Search). Thus, the final PARADISe output shows aside the semi-quantitative report with the areas of the peaks, the Match Factor (MF) of the putative identified analytes, ranging from 0 to 100, which refer to the match with the corresponding mass spectrum in the library.

Only areas of compounds with MF > 80 were considered in this study and the obtained data were organized in a two-dimensional matrix (samples x areas) which was autoscaled and then analysed by PCA.

2.4.4. Software

PCA was carried out by using PLS_Toolbox 8.9.2 software (Eigenvector Research Inc., Manson, WA, USA) for MATLAB®. Designs of experiments were planned with MODDE 9.1 (Umetrics AB, Umeå, Sweden). PARADISe approach was performed by PARADISe software version 3.3. (http://www.models.life.ku.dk/paradise). MLR was carried out using in-house written routines in MATALAB.



Scheme 1. Graphical representation of the initial three-dimensional array having as dimensions: recorded mass fragment (i.e, mass spectra as first mode), elution profiles (second mode) and bread samples (third mode).

3. Results and discussion

3.1. Evaluation of Arrow-SPME and conventional-SPME fiber blanks

As an example, the first acquired blank signal of Arrow-SPME and conventional-SPME fibers, after conditioning step, is reported in Fig. 1A and B, respectively. In both cases, it is quite evident the presence of a high intensity peak in the beginning of the chromatograms (Retention time, R_t:3.70 min), this peak is most likely due to the absorption of atmospheric gas. In addition, some peaks of lower intensity are present along all the signal, most probably originating from the release of compounds belonging to fiber itself. However, Fig. 1A also shows that Arrow-SPME fiber blank is much richer in peaks (R_t: 18–40 min) than the conventional-SPME one, with peaks of higher intensity at R_t: 18 min; R_t: 27 min and R_t: 40 min. In the further analysis the first 7 min of each chromatogram were eliminated because are no informative for the aim of the present study.

Due to the complexity of the Arrow-SPME blank, it was decided to carry out one fiber blank after three sample acquisitions and to compare all the obtained TIC. All the acquired TIC blanks were preprocessed as described in *Section 2.4.2* and were analysed by PCA.

Two components explained 85% of the total variance. The scores plot of the first principal component (PC1), Fig. 1C, shows that blank n° 8 is an evident anomaly, for higher intensity of the whole chromatographic profile (Fig. A5, *Appendix A, Supplementary Materials*). This blank was obtained after a downtime period of acquisition of about two weeks, indicating that the fiber accumulated a series of impurities that could be released during the analysis. The trend of PC2 scores (Fig. 1D) highlights a greater variability between the first blanks (from n ° 1 to n° 10) acquired with the Arrow-SPME and a stabilisation (similar scores values) of the fiber release from the blank signal n° 11 afterwards, thus when the number of experimental runs increases. Therefore, considering the obtained results, it is of paramount importance to clean the fiber before each experimental session and to perform a blank correction before

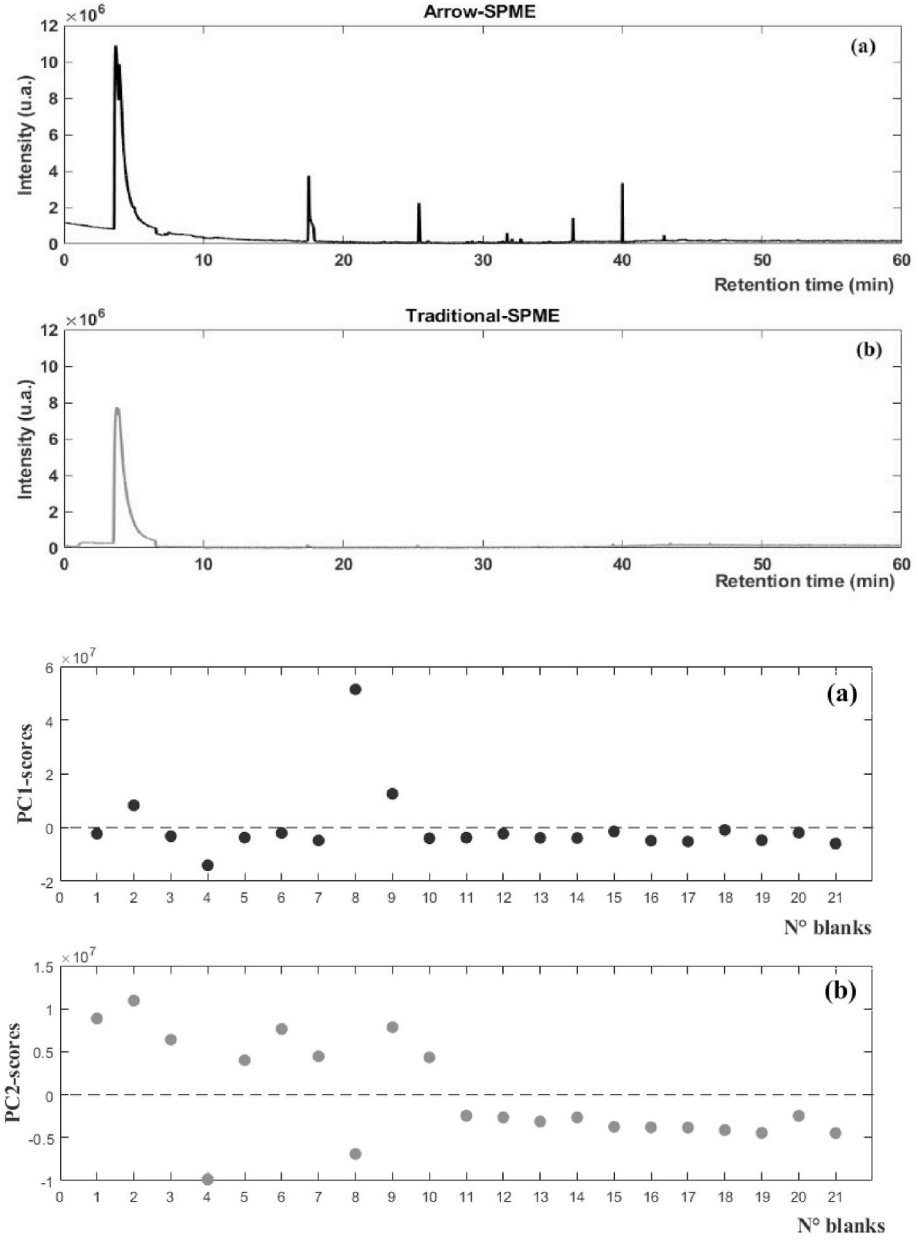


Fig. 1. The first fiber blank acquired after the conditioning of Arrow-SPME (a) and conventional-SPME fiber (b). PC1 (c) and PC2 (d) score plots of PCA performed on ARROW-SPME blank signals.

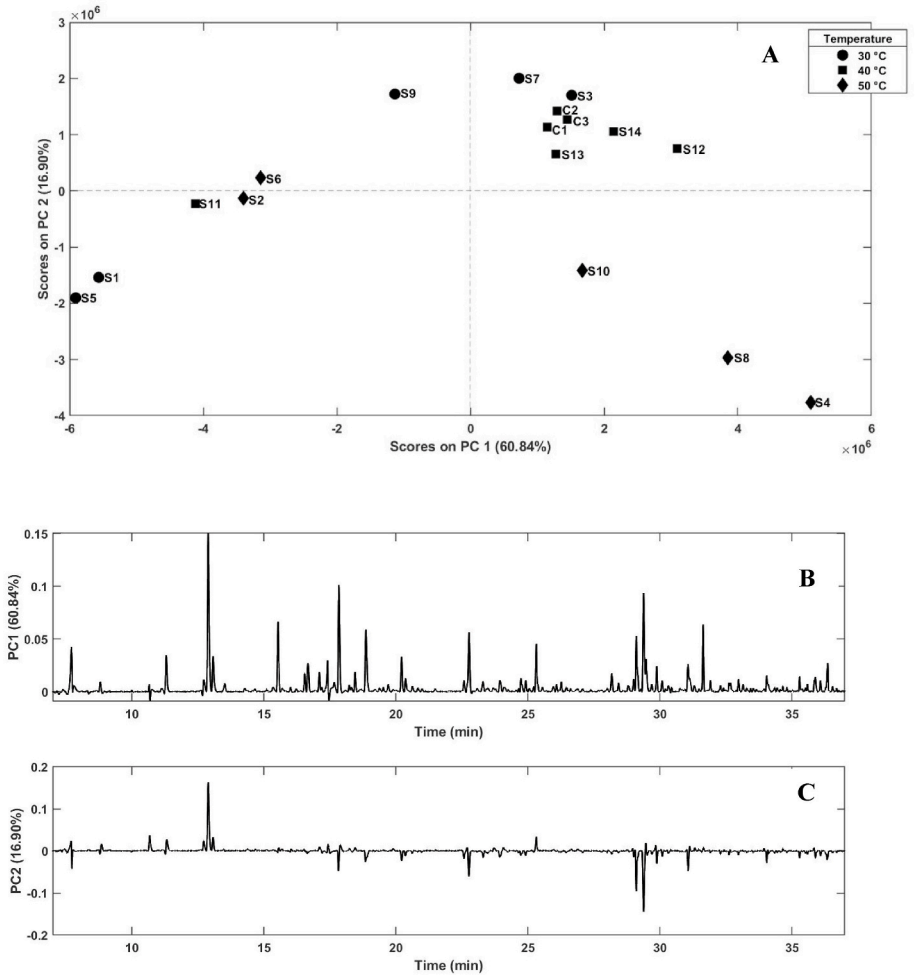


Fig. 2. PC1 *vs.* PC2 scores plot of PCA performed on Arrow-DoE experimental results (A). The experiments with different exposure temperature (T_{exp}) conditions are represented with different symbols. Each DoE-experiment is labelled by its number preceded by 'S', central point replicates by 'C'. PC1 (B) and PC2 (C) loading plot obtained from PCA performed on Arrow-SPME DoE results data.

multivariate analysis of the obtained chromatograms.

3.2. Optimization of SPME extraction conditions through design of experiment

For the optimization of the extraction procedure, all the GC-MS experiments were carried out on the same bread sample, according to the experimental design reported in Table 2.

The optimization and multivariate analysis were separately performed for each fiber. PCA was employed on the chromatographic data to gain valuable insights into the relationships among DoE experimental runs and the analytical responses. The dataset consisted of 17 rows (representing the number of DoE experiments) and 9667 columns (representing TIC chromatographic signals). The chromatograms were pre-processed as explained in *Section 2.4.2*. For the sake of brevity, all the results concerning the Arrow-SPME procedure optimization have been reported in the text, referring the reader to *Appendix B, Supplementary Materials* for the results concerning the conventional-SPME.

The PCA Arrow-SPME model, 2 PCs explaining 78% of total variance, is shown in Fig. 2. Fig. 2B and C reports the loadings plot.

As can be seen from Fig. 2A (PC1 vs. PC2 scores), the three replicates central point runs (C1, C2 and C3) are well grouped and close to the origin of the axes. However, a more detailed discussion about the reproducibility and the performance of the fiber is reported in *Section 3.3*.

Concerning the experiments run in different conditions (DoE experiments S1–S14), it is possible to highlight a clear distinction among some of them. Indeed, PC1 mainly differentiates S8 and S4 (highest positive PC1-scores), at high level of exposure temperature (T_{exp}) and exposure time (see Table 3), from S1 and S5 (highest negative PC1-scores). Furthermore, it is possible to note similarity among samples with different incubation times (S1 and S5, S2 and S6, S3 and S7). Finally, PC2 scores differentiate S3, S7 and S9, at low level of T_{exp} (highest positive PC2-scores), and almost all the runs with medium value for T_{exp} (positive PC2-scores) from all the other experiments.

Inspection of PC1 loadings (Fig. 2A), which accounts for more than 60% of the original variance, evidenced how the contributions of almost all peaks are positive. Notably, these loadings indicate that the scores along the corresponding principal component could be interpreted as an index of the overall recovery of eluted volatile substances [37]. As far as the few negative loadings' values are concerned, they are probably caused by non-perfect Rt alignment. However, in order to better investigate their influence on the interpretability of PC1, the variance accounted for each individual variable on PC1 was investigated (Fig. A6 *Appendix A, Supplementary materials*). Notwithstanding the chromatographic areas with negative loadings have very low intensity (absolute values lower than 0.01), they also present the lowest explained variance with respect to the other chromatographic regions, indicating that PC1 mainly describes the increase in peak area and not the peak shift. Finally, the differences among PC2 seem to be mainly due to 1-butanol, 3 – methyl (R_t: 13.05 min) with positive loading values and to nonanal (R_t: 29.29 min) and phenylethyl alcohol (R_t: 29.62 min) with negative

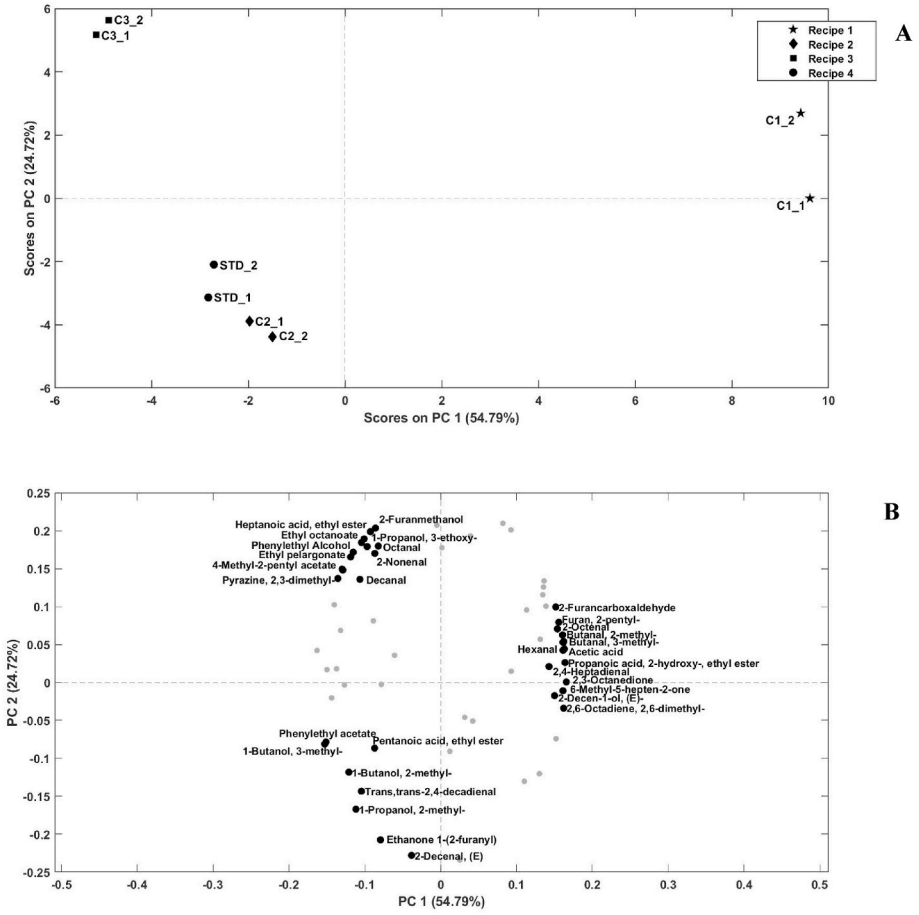


Fig. 3. Analysis of VOCs of bread samples sampled with Arrow-SPME. PC1 vs PC2 scores plot obtained by PCA applied on selected areas from PARADISe analysis (A). Analysis of VOCs of bread samples sampled with Arrow-SPME. PC1 vs PC2 loadings plot obtained by PCA applied on selected areas from PARADISe analysis (B).

loading values (Fig. 2B).

Since, the difference observed along PC1 in the scores plot can be attributable to the intensity of the entire chromatographic profile (almost all loadings are positive), the experimental settings corresponding to S4 and S8 runs (Table 2b) seem to give better performance with respect to the other runs setting. Considering the results shown in Fig. 2B, the first principal component well represents the entire chromatographic profile and thus the PC1 scores values were used as a response to be optimised. This approach made it possible to simplify the complexity of the analysis, since PC1 summarises the main variance of the data [37].

After this preliminary exploratory analysis, a multilinear regression model (MLR) was built to obtain a complete and exhaustive treatment of the data and to develop a predictive model for the chosen response. The applied experimental design allowed the estimation of the coefficients (b) of the following postulated model:

$$\mathbf{y} = b_0 + b_T\mathbf{T_{exp}} + b_{t.exp}\mathbf{t_{exp}} + b_{t.inc}\mathbf{t_{inc}} + b_T^2\mathbf{T_{exp}^2} + b_{t.inc}^2\mathbf{t_{inc}^2} + b_{t.exp}^2\mathbf{t_{exp}^2} + b_{Texp\_tinc}\mathbf{T_{exp}t_{inc}} + b_{Texptexp}\mathbf{T_{exp}t_{exp}} + b_{texptinc}\mathbf{t_{exp}t_{inc}} \quad \text{(Equ.3)}$$

Where $\mathbf{y}$ is the PC1 scores vector and T_{exp}, t_{exp} and t_{inc} are the values of exposure temperature and time, and incubation time, respectively. The obtained results are summarised in Table A1 *Appendix A*, where the values of the significant regression coefficients, b_i, were reported. Significance of each term was computed at a significance level of 5%. As far as the main terms are concerned, both exposure time and temperature are significant and positive, and they indicate that the best VOCs sampling condition is obtained when high values of these factors are considered. All the other coefficients have no significant effect except for the squared term $(t_{exp})^2$, which implies that the relation between the factor and the response is not linear. Its negative term could have a physical explanation due to the sampling limited volume of the stationary phase able to adsorb/absorb analytes, beyond which no gains in terms of sampled analytes can be achieved.

Considering the non-linearity of the relation between the t_{exp} factor and the response, t_{exp} and T_{exp} values, corresponding to the optimal separation conditions has been investigated by means of a contour plot (Fig. A7, *Appendix A, Supplementary Materials*), where it was possible to note that the maximum of the response is achieved assuming both these factors at their high values.

In summary, the analysis of DoE results indicates as the best operative conditions for the extraction and sampling of VOCs by the Arrow-SPME fiber: high level for T_{exp} (50 °C) and t_{exp} (30 min). The t_{inc} since had no significant effect on the final response, was kept low (10 min) to reduce analysis time.

The same optimal experimental conditions were obtained for extraction with the conventional- SPME fiber, as reported in Appendix B, Supplementary Materials.

3.3. Comparison of arrow and conventional SPME performance

To test the reproducibility of both fibers, the chromatographic signals coming from the analysis of the control sample were considered. As mentioned in *Session 2.1*, the control sample was a store-bought saltines sample and was independently analysed five times with both fibers under their respective optimised conditions. The corresponding signals were analysed by means of the PARADISe approach (as better explained in *Sections 2.4.3 and 3.2*); sixteen compounds were identified with a MF $>$ 80 and areas and RSD values corresponding to the individuated peak

extracted from Arrow-SPME and conventional-SPME fibers are reported in Table 5.

From the obtained data is possible to assess that significantly higher sensitivity, compared to conventional-SPME fiber, is achieved by the SPME-Arrow, even reaching differences of up to one order of magnitude. This enhanced sensitivity can be surely attributed to the Arrow-SPME's larger sorptive phase and expanded surface area, enabling superior sample capturing capabilities. Despite some cases, where the relative standard deviation (RSD) exceeded 10%, the Arrow-SPME consistently produced more reproducible results as well.

3.4. VOCs compounds determination and identification by PARADISe

Each bread sample was twice measured, and all TIC chromatograms were pretreated as described in *Section 2.4.2* and arranged in a three-dimensional array of 396 × 9617 x 8 dimensions.

The defined intervals were 130, selected in order to enclose, when possible, only a single peak. In PARADISe the number of explored factors, in the PARAFAC2 modelling, ranged from 1 to 7 and the maximum number of iterations to reach convergence was set equal to 3000 (default parameter). The mass spectra libraries quested were NIST 08 and Wiley 275.

The obtained model, for each interval, showed resolved chromatographic profiles, as well as baseline effects, some interfering effects, and low signal-to-noise peaks. Given the complexity of the product, the choice of the right number of components could be a critical aspects, however, in PARADISe, it is supported by a graphical interface, showing the results obtained by a deep learning routine trained to recognize "a peak" from "not a peak" [33,38]. In this way the user is guided in the selection of the components corresponding to the chemical information (peaks). For the sake of clarity, in Fig. A8, *Appendix A, Supplementary Materials*, it is reported an example of the graphical interface which summarised the information used for modelling one of the selected intervals (R_t from 16.5 min to 17.2 min) considered in this study.

The individuated analytes were refined by eliminating a series of aliphatic and aromatic hydrocarbons which, as reported in the literature, do not contribute significantly to the aroma [11], obtaining 66 final compounds (see Table A2 *in Appendix A, Supplementary materials*).

3.5. PCA investigation

The obtained areas were organised in a two-dimensional matrix (8 bread samples *x* 66 areas); the data were autoscaled and analysed by PCA analysis, three principal components (explaining 90% of total variance) are discussed.

Table 5
VOCs area and relative standard deviation (RSD%) obtained from the analysis of control sample.

Control sample analysis	Arrow		Conventional	
	Area	RSD (%)	Area	RSD (%)
Acetic acid	2410230	9	461945	16
Butanal, 3-methyl	365913	23	34635	36
Butanal, 2-methyl	142184	22	11372	42
Hexanal	1456642	10	304467	32
Pyrazine, methyl	35258	24	20791	60
Iso-amyl acetate	29500	13	11081	22
Benzene, 1,4-dimethyl	37027	17	21276	32
Heptanal	79576	5	56724	8
2-Heptenal	21796	9	11594	27
Benzaldehyde	75167	11	51002	16
Furan, 2-pentyl	153384	6	95866	18
N-octanal	52889	9	70849	13
2-octenal	15956	12	12499	18
Octanoic acid, ethyl ester	28454	25	21023	13
Decanal	64840	13	65471	52
Dodecane, 2,6,10-trimethyl-	15876	19	14471	17

From the scores plot of the first two components (Fig. 3A), it is possible to highlight differences between samples as a function of their different recipes. In particular, the first principal component clearly distinguishes C1 (bread obtained with sourdough) samples from others. Instead, according to PC2, C3 (bread with brewer's yeast at 3 h of fermentation time) was on the positive side, while C2 (bread with a mixture of yeast at 1 h of fermentation time) and STD (bread with brewer's yeast at 1 h of fermentation time) on the negative one. The highlighted distribution of investigated bread samples in the score's domain, indicate that variation of VOCs profile of bread is surely dependent by used yeast as well as by the leavening time. Indeed, from Fig. 3B, C1 samples, located on the positive side of PC1, seem to be richer in aldehydes, carboxylic acids and esters, namely acetic acid, 2-methyl-butanal, 3-methyl-butanal, hexanal, 2-pentyl-furan, 2-furancarboxaldehyde (furfural), 2-octenal, 2-hydroxy-propanoic acid ethyl ester (lactic acid ethyl ester), (trans, trans)-2,4-heptadienal. There are also some ketones (2,3-octanedione, 6-methyl-5-hepten-2-one) and others compounds such as 2-decen-1-ol and 2,6-dimethyl-2,6-octadiene that differentiate C1 from the other bread samples. This chemical composition was also observed in other studies [11,17], in fact, acetic acid is one of the most important products derived from bacteria present in sourdough and its content is fundamental in the characterization of the aroma of bread [39]. Among the different individuated compounds, 2-methyl-butanal is another fermentation compound that probably comes from the so-called "Ehrlich pathway" in yeast cells by leucine and iso-leucine [12] while 3-methyl-butanal is part of the so-called "Strecker's aldehydes" which result from the conversion of amino acids in yeast metabolism. As regards 3-methyl-butanal, it could be derived from the conversion of isoleucine [14] and this compound could also give bread a strong malty aroma. For the sake of clarity, in Table A2 (*Appendix A, Supplementary materials*), the characteristic aroma and Odour Threshold value (OTV) in water for each individuated compound, found in literatures [19,40,41], were reported, when possible. Hexanal and 2-pentylfuran are the main products of the oxidation of lipids contained in flour that occurs during the metabolism of yeasts. Hexanal is typical of the aroma of breadcrumbs, and it is reported as an "unpleasant odour", along with other compounds that always result from lipid oxidation, such as nonanal. Consequently, a high concentration of this compound may result in a less approval from sensory analysis [12,19] As concerns 2-pentylfuran, since the type of flour used for the samples analysed in this study is the same in all samples, its presence could be strictly related to the used high fermentation time of the dough used for the production of C1 bread (5 h). Finally, in this sample is also present furfural compound, obtained by caramelization and non-enzymatic Maillard reactions [16], which mainly characterises the aroma of bread giving the typical smell of bread crust [19] and the (trans, trans)-2,4-heptadienal which is reported to be one of the major contributors to the rancid smell of used olive oil [42].

The differences attributable to C3 samples could be mainly due to a series of aldehydes and alcohols. Thereby from the literature, it is known that, octanal and the other aldehydes derive from the lipid oxidation of yeast [12,43] and the ethyl ester of heptanoic acid probably comes from enzymatic reactions that take place inside yeast cells [19]; 2-phenylethyl alcohol is derived from the catabolism of phenylalanine in the Erhlich pathway [14]. Furthermore, C3 samples seem to be also rich in esters and pyrazine derivative, namely heptanoic acid ethyl ester, ethyl pelargonate, ethyl octanoate, 4-methyl-2-pentyl acetate and 2, 3-dimethyl-pirazine.

C2 and STD samples are mainly characterised by the following analytes: phenylethyl acetate, 3-methyl-1-butanol, 2-methyl-1-butanol, pentanoic acid ethyl ester, (trans, trans)-2,4-decadienal, 2-methyl-1-propanol, 1-(2-furanyl) ethanone, 2-decenal (E). 3-Methyl-1-butanol can be produced after typical sourdough fermentation [15]. The presence of alcohols such as 2-methyl-1-propanol is attributable to the reduction process of low molecular weight aldehydes, such as acetaldehyde, by enzymes such as dehydrogenase [44].

PC3 explained 11% of data variability and mainly differentiate STD samples (negative score values) from C2 ones (positive scores value) mainly to the presence of some of carboxylic acids and their derivatives (positive PC3-scores) and some of the identified alcohol compounds (negative PC3-scores) (Fig. A9 and A10, *Appendix A, Supplementary Materials*).

4. Conclusions

In this study, the performance of SPME Arrow was compared with a conventional SPME fiber constituted by the same chemical stationary phase, but with different physical characteristics, i.e. higher amount of stationary phase and major diameter. In particular, the main parameters affecting the performance of the SPME Arrow and conventional fibers in sampling VOCs of bread samples were investigated and optimised by means an untargeted approach based on the synergistic use of Experimental Design tool and multivariate data analysis. The same sample (saltines samples) was analysed different times with both the fibers and their performances were evaluated in terms of reproducibility and sensitivity. The use of the SPME Arrow fiber allowed to obtain the best results and these aspects were considered an important prerequisite able to better investigate the influence of the different type of yeasts and fermentation time on VOCs profile of the bread.

Therefore, the optimised condition of SPME Arrow together with GC-MS measurements and a chemometrics approach were used for the chemical characterization of aroma profile of several industrial bread samples obtained with different receipt. In particular, the use of PARAFAC2 automated in PARADISe allowed the determination and the identification of 66 compounds. The identified compounds mainly come from the fermentation reaction of bacteria, caramelization and non-enzymatic Maillard reactions, yeast metabolism and reaction and lipid oxidation of flour. Thanks to the use of this analytical method, it was possible to highlight differences in bread samples as function of individuated volatile organic compounds, which seemed to be more accentuated in samples obtained by sourdough at 5 h of fermentation. On the other hand, having used a mixture of sourdough and brewer's yeast did not produce a higher difference in aroma profile with respect to bread samples obtained only with brewer's yeast. This similarity was probably due to the use of the same leaving time (1 h) showing a great influence of the latter parameter on the development of final VOCs profile as well. Furthermore, among the different individuated compounds, some of them were not surely correlated with a pleasant aroma associated with breads. It is worth to mention that this aspect is of utmost of importance, since the presence of analytes that, also in low concentration, are characterised by high odour intensities could make the difference between a "desired" or "undesired" aroma by the consumer. In fact, considering all the results obtained in this study, on one side, it can be concluded that C1 samples, obtained with sourdough at 5 h of leaving time had a more complex VOCs profile, but, on the other side, their differences are also due to the presence of minority compounds that could negatively influence their aroma. Therefore, having an analysis tool that allows to highlight even these small differences can be of paramount importance in the characterization and evaluation of new food formulation.

Consent for publication

Written informed consent for publication was obtained from all participants.

CRediT authorship contribution statement

S.Pellacani: investigation, methodology, software, data curation; formal analysis, writing; C. Durante: conceptualization, methodology, project administration, formal analysis, software, supervision, writing – original draft preparation; S.Celli, M.Mariani: resources, investigation; A. Marchetti: investigation, methodology, writing – original draft preparation. L. Strani: investigation, methodology. M. Cocchi: software, methodology, formal analysis and writing – original draft preparation, writing—review and editing.

Novelty statement

The developed analytical method is based on the synergistic use of an Arrow fiber for sampling volatile organic compounds and a deep learning (PARADISe) approach for the extraction of useful information from signals obtained by the analysis through GC-MS method. The developed analytical method was used to study the flavour of four industrial bread samples and to deepen and to test the impact of different types of yeast and their respective leaving time on the VOCs produced in the final product.

Since Arrow fiber is a fairly novel device, its performance was thoroughly investigated in terms of blanks, reproducibility and sensitivity in comparison to a traditional SPME fiber. The inspection of SPME-Arrow blank is never discussed in previous studies and information about all the three investigated aspects is of utmost of relevance in the study of food aroma, since this device, based on a novel SPME geometry, can be used to "hunt" target molecules more efficiently in complex matrices.

Furthermore, PARADISe approach resulted to be of utmost importance since it has been found to be an extremely useful tool for the deconvolution of the chromatographic peaks directly from raw chromatographic data and simultaneous integration of the areas of the deconvoluted peaks for all samples. Finally, Principal Component Analysis was performed on the peak areas resolved by PARADISe to obtain preliminary information on the main VOCs present in the studied bread samples, taking also into consideration the interplay of yeast and leavening time.

To the best of the author's knowledge, the bread aroma has never been sampled and analysed using the proposed analytical approach. Furthermore, this method allows for achieving a more precise understanding of the aroma and the parameters that influence its variation. This aspect is of paramount importance for the baking industry to adjust recipe modifications and production parameters, as well as to meet consumer needs in formulating new products.

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

The data that has been used is confidential.

Acknowledgement

This research was carried out thanks to FAR2022 project of Department of Chemical and Geological Science, University of Modena and Reggio Emilia, Italy.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chemolab.2023.104940.

References

[1] H. Wang, P. Yang, C. Liu, H. Song, W. Pan, L. Gong, Characterization of key odor-active compounds in thermal reaction beef flavoring by SGC×GC-O-MS, AEDA, DHDA, OAV and quantitative measurements, J. Food Compos. Anal. 114 (2022), https://doi.org/10.1016/j.jfca.2022.104805.

[2] M. Starowicz, Analysis of volatiles in food products, Separations 8 (2021) 157, https://doi.org/10.3390/separations8090157.
[3] Y.-H. Huang, T.-H. Kao, B.-H. Chen, Development of a GC-MS/MS method coupled with HS-SPME-Arrow for studying formation of furan and 10 derivatives in model systems and commercial foods, Food Chem. 395 (2022), 133572, https://doi.org/10.1016/j.foodchem.2022.133572.
[4] X.-B. Xu, K. Murtada, J. Pawliszyn, Determination of selected volatile terpenes in fish samples via solid phase microextraction arrow coupled with GC-MS, Talanta 221 (2021), 121446, https://doi.org/10.1016/j.talanta.2020.121446.
[5] T.G. Nam, J.-Y. Lee, B.-K. Kim, N.-E. Song, H.W. Jang, Analyzing volatiles in brown rice vinegar by headspace solid-phase microextraction (SPME)–Arrow: optimizing the extraction conditions and comparisons with conventional SPME, Int. J. Food Prop. 22 (2019) 1195–1204, https://doi.org/10.1080/10942912.2019.1634099.
[6] N. Manousi, E. Rosenberg, G.A. Zachariadis, Solid-phase microextraction arrow for the sampling of volatile organic compounds in milk samples, Separations 7 (2020) 75, https://doi.org/10.3390/separations7040075.
[7] I. Šikuten, P. Štambuk, J. Karoglan Kontić, E. Maletić, I. Tomaz, D. Preiner, Optimization of SPME-arrow-GC/MS method for determination of free and bound volatile organic compounds from grape skins, Molecules 26 (2021) 7409, https://doi.org/10.3390/molecules26237409.
[8] X. Zhang, C. Wang, L. Wang, S. Chen, Y. Xu, Optimization and validation of a head space solid-phase microextraction-arrow gas chromatography-mass spectrometry method using central composite design for determination of aroma compounds in Chinese liquor (Baijiu), J. Chromatogr., A 1610 (2020), 460584, https://doi.org/10.1016/j.chroma.2019.460584.
[9] Q. Pan, J. Zhou, W. Shen, Z. Wang, H. Cai, X. Jia, Effect of extruded wheat bran on volatile and physicochemical properties of bread during its shelf life, J. Cereal. Sci. 107 (2022), 103525, https://doi.org/10.1016/j.jcs.2022.103525.
[10] L.S. Amann, O. Frank, C. Dawid, T.F. Hofmann, The sensory-directed elucidation of the key tastants and odorants in sourdough bread crumb, Foods (2022) 11, https://doi.org/10.3390/foods11152325.
[11] L.M. Seitz, O.K. Chung, R. Rengarajan, Volatiles in selected commercial breads, Cereal Chem. 75 (1998) 847–853, https://doi.org/10.1094/CCHEM.1998.75.6.847.
[12] A.N. Birch, M.A. Petersen, N. Arneborg, Å.S. Hansen, Influence of commercial baker's yeasts on bread aroma profiles, Food Res. Int. 52 (2013) 160–166, https://doi.org/10.1016/j.foodres.2013.03.011.
[13] M.S. Zolfaghari, S.M.S. Ardebili, G.H. Asadi, K. Larijani, Effect of sourdough, bakery yeast and sodium bicarbonate on volatile compounds, and sensory evaluation of lavash bread, J. Food Process. Preserv. 41 (2017), e12973, https://doi.org/10.1111/jfpp.12973.
[14] E. Aslankoohi, B. Herrera-Malaver, M.N. Rezaei, J. Steensels, C.M. Courtin, K. J. Verstrepen, Non-conventional yeast strains increase the aroma complexity of bread, PLoS One 11 (2016), e0165126, https://doi.org/10.1371/journal.pone.0165126.
[15] R. Niçin, N. Özdemir, Ö. Şimşek, A.H. Çon, Production of volatiles relation to bread aroma in flour-based fermentation with yeast, Food Chem. 378 (2022), 132125, https://doi.org/10.1016/j.foodchem.2022.132125.
[16] L. De Luca, A. Aiello, F. Pizzolongo, G. Blaiotta, M. Aponte, R. Romano, Volatile organic compounds in breads prepared with different sourdoughs, Appl. Sci. 11 (2021) 1330, https://doi.org/10.3390/app11031330.
[17] I.H. Cho, D.G. Peterson, Chemistry of bread aroma: a review, Food Sci. Biotechnol. 19 (2010) 575–582, https://doi.org/10.1007/s10068-010-0081-3.
[18] Salim-ur-Rehman, A. Paterson, J.R. Piggott, Flavour in sourdough breads: a review, Trends Food Sci. Technol. 17 (2006) 557–566, https://doi.org/10.1016/j.tifs.2006.03.006.
[19] A.N. Birch, M.A. Petersen, Å.S. Hansen, REVIEW: aroma of wheat bread crumb, Cereal Chem. 91 (2014) 105–114, https://doi.org/10.1094/CCHEM-06-13-0121-RW.
[20] E.C. Garvey, M.G. O'Sullivan, J.P. Kerry, K.N. Kilcawley, Optimisation of HS-SPME parameters for the analysis of volatile compounds in baked confectionery products, Food Anal. Methods 13 (2020) 1314–1327, https://doi.org/10.1007/s12161-020-01740-4.
[21] N. Tungkijanansin, W. Alahmad, T. Nhujak, P. Varanusupakul, Simultaneous determination of benzoic acid, sorbic acid, and propionic acid in fermented food by headspace solid-phase microextraction followed by GC-FID, Food Chem. 329 (2020), 127161, https://doi.org/10.1016/j.foodchem.2020.127161.
[22] C. Lancioni, C. Castells, R. Candal, M. Tascon, Headspace solid-phase microextraction: fundamentals and recent advances, Adv. Sample Prep. 3 (2022), 100035, https://doi.org/10.1016/j.sampre.2022.100035.
[23] A. Spietelun, A. Kloskowski, W. Chrzanowski, J. Namieśnik, Understanding solid-phase microextraction: key factors influencing the extraction process and trends in improving the technique, Chem. Rev. 113 (2013) 1667–1685, https://doi.org/10.1021/cr300148j.
[24] L.G. Johnsen, P.B. Skou, B. Khakimov, R. Bro, Gas chromatography – mass spectrometry data processing made easy, J. Chromatogr., A 1503 (2017) 57–64, https://doi.org/10.1016/j.chroma.2017.04.052.
[25] E.C. Garvey, M.G. O'Sullivan, J.P. Kerry, K.N. Kilcawley, Optimisation of HS-SPME parameters for the analysis of volatile compounds in baked confectionery products, Food Anal. Methods 13 (2020) 1314–1327, https://doi.org/10.1007/s12161-020-01740-4.
[26] J. Herrington, G. Gómez-Ríos, C. Myers, G. Stidsen, D. Bell, Hunting molecules in complex matrices with SPME arrows: a review, Separations 7 (2020) 12, https://doi.org/10.3390/separations7010012.
[27] R. Leardi, Experimental design in chemistry: a tutorial, Anal. Chim. Acta 652 (2009) 161–172, https://doi.org/10.1016/j.aca.2009.06.015.
[28] T. Lundstedt, E. Seifert, L. Abramo, B. Thelin, Å. Nyström, J. Pettersen, R. Bergman, Experimental design and optimization, Chemometr. Intell. Lab. Syst. 42 (1998) 3–40, https://doi.org/10.1016/S0169-7439(98)00065-3.
[29] M. Li Vigni, C. Durante, M. Cocchi, Chapter 3 - exploratory data analysis, in: F. Marini (Ed.), Data Handl. Sci. Technol., Elsevier, 2013, pp. 55–126, https://doi.org/10.1016/B978-0-444-59528-7.00003-X.
[30] G. Tomasi, F. Savorani, S.B. Engelsen, icoshift: an effective tool for the alignment of chromatographic data, J. Chromatogr., A 1218 (2011) 7832–7840, https://doi.org/10.1016/j.chroma.2011.08.086.
[31] P. Eilers, H. Boelens, Baseline Correction with Asymmetric Least Squares Smoothing, Unpubl Manuscr., 2005.
[32] J.M. Amigo, T. Skov, R. Bro, J. Coello, S. Maspoch, Solving GC-MS problems with PARAFAC2, TrAC, Trends Anal. Chem. 27 (2008) 714–725, https://doi.org/10.1016/j.trac.2008.05.011.
[33] G. Baccolo, B. Quintanilla-Casas, S. Vichi, D. Augustijn, R. Bro, From untargeted chemical profiling to peak tables – a fully automated AI driven approach to untargeted GC-MS, TrAC, Trends Anal. Chem. 145 (2021), 116451, https://doi.org/10.1016/j.trac.2021.116451.
[34] R. Bro, PARAFAC. Tutorial and applications, Chemometr. Intell. Lab. Syst. 38 (1997) 149–171, https://doi.org/10.1016/S0169-7439(97)00032-4.
[35] M.H. Kamstrup-Nielsen, L.G. Johnsen, R. Bro, Core consistency diagnostic in PARAFAC2, J. Chemom. 27 (2013) 99–105, https://doi.org/10.1002/cem.2497.
[36] Tutorial on PARADISe: PARAFAC2-Based Deconvolution and Identification System for Processing GC–MS Data, 2023, https://doi.org/10.21203/rs.3.pex-2143/v1.
[37] V. Giannetti, M. Boccacci Mariani, F. Marini, A. Biancolillo, Effects of thermal treatments on durum wheat pasta flavour during production process: a modelling approach to provide added-value to pasta dried at low temperatures, Talanta 225 (2021), 121955, https://doi.org/10.1016/j.talanta.2020.121955.
[38] A.B. Risum, R. Bro, Using deep learning to evaluate peaks in chromatographic data, Talanta 204 (2019) 255–260, https://doi.org/10.1016/j.talanta.2019.05.053.
[39] S. Wu, Y. Peng, J. Xi, Q. Zhao, D. Xu, Z. Jin, X. Xu, Effect of sourdough fermented with corn oil and lactic acid bacteria on bread flavor, LWT 155 (2022), https://doi.org/10.1016/j.lwt.2021.112935.
[40] K.R. Cadwallader, Aromas, in: Encycl. Food Chem., Elsevier, 2018, pp. 22–29, https://doi.org/10.1016/B978-0-08-100596-5.21623-5.
[41] F. Shahidi, A. Abad, Lipid-Derived Flavours and Off-Flavours in Food, 2018, https://doi.org/10.1016/B978-0-08-100596-5.21666-1.
[42] H. Zhu, S.C. Wang, C.F. Shoemaker, Volatile constituents in sensory defective virgin olive oils, Flavour Fragrance J. 31 (2016) 22–30, https://doi.org/10.1002/ffj.3264.
[43] J. Pico, J. Bernal, M. Gómez, Wheat bread aroma compounds in crumb and crust: a review, Food Res. Int. 75 (2015) 200–215, https://doi.org/10.1016/j.foodres.2015.05.051.
[44] L.-A. Payet, M. Leroux, J.C. Willison, A. Kihara, L. Pelosi, F. Pierrel, Mechanistic details of early steps in coenzyme Q biosynthesis pathway in yeast, Cell Chem. Biol. 23 (2016) 1241–1250, https://doi.org/10.1016/j.chembiol.2016.08.008.