Contents lists available at ScienceDirect

Talanta

journal homepage: http://ees.elsevier.com

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

Vanessa Giannetti [a,*], Maurizio Boccacci Mariani [a], Federico Marini [b], Alessandra Biancolillo [c]

[a] Department of Management, Sapienza University of Rome, Via Del Castro Laurenziano 9, 00161, Rome, Italy
[b] Department of Chemistry, Sapienza University of Rome, P.le Aldo Moro 5, 00185, Rome, Italy
[c] Department of Physical and Chemical Sciences, University of L'Aquila, Via Vetoio, 67100, Coppito, L'Aquila, Italy

ARTICLE INFO

Keywords
Pasta
Headspace solid phase microextraction (HS-SPME)
Design of experiments (DoE)
Drying process
Partial least squares discriminant analysis (PLS-DA)
High/very high temperature-short time (HT/VHT-St) and low temperature-long time (LT-Lt)

ABSTRACT

Pasta is a key element of the Mediterranean Diet and it has been declared by Unesco *intangible cultural heritage of humanity*. Despite seems a simple food, only made of semolina and water, pasta is produced following a multi-step process that strongly affect the final product. Drying stage is the one that has the greater influence on its organoleptic/nutritional characteristics. This study aimed to analyse the flavour of pasta to test whether the different drying treatments (*High Temperature-Short time* or *Low Temperature-Long time*) have a direct impact on its composition and consequently whether they could influence the end-product quality. The headspace solid-phase microextraction was optimized using an experimental design and 52 samples were analysed by HS-SPME/GC-MS and classified by PLS-DA. The resulting classification model (validated by repeated double cross-validation and permutation tests) allowed correctly predicting more than 80% of samples, confirming that drying may have a significant impact on pasta flavour.

1. Introduction

Pasta is obtained by extruding, rolling, drying and packaging of dough prepared exclusively with wheat semolina and water, without any addition of colouring agents or preservatives. Pasta is a versatile, functional, high genuineness, and long shelf-life product, but the reason of its success and spread worldwide also lies in the simplicity of the technological steps in its production and its extremely simple composition (flour and water), that can be easily adjusted substituting cheaper and easier to trace cereals to encounter any cultural needs and food traditions.

Pasta is a key element of the Mediterranean Diet. WHO (World Health Organization of the United Nations) and FAO (Food and Agriculture Organization of the United Nations) described pasta a *healthy, sustainable and quality food model*, and in 2010, UNESCO (United Nations Educational, Scientific and Cultural Organization) declared it an *intangible cultural heritage of humanity*.

Pasta plays an important role in the Italian food tradition and with over 55% exports contributes to promote and disseminate its economic and cultural heritage value of the made-in-Italy food around the world (Aidepi - *Italian association of confectionery and pasta industries*, 2017). According to Coldiretti (*National federation of Italian farmers*), Italian consumers prefer dry pasta, representing about 90% of the total purchasing volume of the sector. From the supply point of view, the domestic market is highly fragmented, only one player with a stable market share around 32%, followed by some national competitors with market shares far below (12.3%, 8.1% and 6.5%; 13.5% private labels; and 27.5% other producers) and around 600 small and very small-sized manufacturing factories located nation-wide (IRI, 2014). In order to preserve its historical identity, the Italian Law No. 580 of 1967 (amended by DPR No. 187/2001) was issued to well-define the requirements concerning the qualitative characteristics of raw materials imposing the exclusive use of durum wheat semolina in its production, but avoiding to mention restrictions concerning production technologies [1]. Conditions in the manufacturing process can significantly affect semolina components, threatening the competition between starch and protein during cooking, and influencing the aromatic profile of end-product [2–4]. In the manufacturing process, the drying step is certainly the stage that has the greater influence on the final product peculiarities: it can be LT-Lt (*Low Temperature-Long time*) or HT/VHT-St (*High/Very High Temperature-Short time*). The temperature varies considerably according to the drying technology used: < 60 °C for drying at low temperatures (LT), 70–80 °C for drying at high temperatures (HT) until at 110 °C for that at very high temperature (VHT). Consequently, the drying period also varies, longer than 20 h for long-time drying treatments (LT-Lt), about 7–10 h for short-time processes (HT-St) and until 1–2 h for very short-time treatments (VHT-St). Over the years, the

* Corresponding author.
E-mail address: vanessa.giannetti@uniroma1.it (V. Giannetti)

https://doi.org/10.1016/j.talanta.2020.121955
Received 14 September 2020; Received in revised form 23 November 2020; Accepted 30 November 2020
Available online xxx
0039-9140/

technological progress and the constant increase of higher hygiene standards, together with a change of market demand, have gradually modified the traditional pasta low-temperature processing method. Most producers, therefore, have oriented their production towards HT/VHT-St technologies. The advantages of these technologies are considerable, such as an increase of production due to a reduction on production times, product shelf-life extension, or texture property improvements during cooking [5,6]. On the other hand, several studies showed that the extreme thermal stress, in the drying step, negatively affect both nutritional (loss in essential amino acids with reduction of nutritional value) and organoleptic (formation of Maillard reaction by-products resulting in a flattening of flavour and browning of colour) pasta properties while low drying temperatures (≤65 °C), limiting the heat-damage, better preserve the characteristics of the final product [7–10]. In literature, Maillard reaction by-products, lipid oxidation by-products, neo-formation aromatic compounds, microbiological processes by-products (fermentation processes) and the compounds that cause a reduction in nutritional value, were all used as indicators to describe the influence of the manufacturing process on food quality [11–16].

The aim of our studies is to analyse the flavour profile of made-in-Italy pasta to prove that different drying treatment (LT-Lt, HT or VHT-St) have a direct impact on its qualitative/quantitative composition and, consequently, it could influence the nutritional and organoleptic quality of the end-product.

Among pasta's main quality parameters, measured by standard procedures, such as colour, cooking resistance, and texture properties, further information on the characteristics and *history* of the product can also be obtained investigating other molecular markers. Considering the conventional indicators, routine controls often neglect technological and organoleptic aspects that significantly contribute its total quality [17]. Although these aspects are not perceived directly at the moment of consumption, they represent an added value to be considered in order to discriminate *premium price* products. In this context, flavour could be used to valorise pasta that retains all its features at the end of the production process.

In a previous study, published in Boccacci Mariani et al. [18], our team developed an HS-SPME/GC-MS (headspace solid-phase microextraction/gas chromatography-mass spectrometry) screening method to investigate the volatile fraction on durum wheat dry pasta samples. That method adopted: i) the traditional analytical approach in order to optimize those parameters that compromise extraction efficiency of HS-SPME technique (i.e. the selection of the fibre coating, the influence of sampling temperature and time, the effect of the sample weight); ii) the use of standard solutions to optimize the HS-SPME procedure and create a flavour-compounds library in order to identify possible discriminant aromatic substances in pasta samples; iii) finally, the use of a gas-chromatograph interfaced with an ion-trap mass spectrometer to monitor sample chromatographic profiles. In order to verify the feasibility of this method, a further study was later carried out by our team that analysed a larger number of pasta samples, as published in Giannetti et al. [19]. The sample-set consisted of pasta belonging to two different classes, i.e. pasta obtained with LT-Lt drying treatment (*traditional products*) and pasta produced using HT/VHT-St drying step (*industrial products*). The results obtained by principal component analysis (PCA) have shown that a well-defined separation among the two groups of the considered pasta samples is possible.

Based on these studies, confirmed by others available in literature, it is thus possible to characterize pasta samples in relation to the drying conditions used for its production [20]. Therefore, the flavour may help to discriminate pasta samples manufactured with the traditional method (LT-Lt) from those obtained under extreme drying conditions.

The successful results provided by our previous work have encouraged us to further develop both the HS-SPME technique and chemometric models to identify a set of volatiles compounds that could be used as product or process markers. Indeed, although some volatile compounds found in the end product are already present in the aromatic profile of durum wheat (mainly alcohols, esters, aldehydes, terpenes), many other, e.g. heterocyclic compounds (furans, pyrroles, pyrazines, thiazoles) and volatiles carboxylic acids, are formed during drying step in relation to the process conditions adopted (temperature, time, moisture, static or dynamic cell) [21,22]. Those markers could help to assess the effects of industrial production technologies on the final characteristics of the pasta samples belonging to two specific market segments (*traditional* or *industrial* products), to keep the drying step under control or to assess product authenticity compared to what mentioned in the label. In this work, the pasta-flavour monitoring was carried out developing an HS-SPME/GC-MS method, whose extraction conditions were optimized using a rational design of experiments and, in particular, a response surface methodology.

Eventually, Partial Least Squares Discriminant Analysis (PLS-DA) has been used to discriminate pasta produced under HT/VHT-St and LT-Lt drying process. This classifier has been chosen because it has been widely and successfully exploited in the food analysis context and, in particular, for quality control of pasta. In fact, it has been used to authenticate PGI Gragnano pasta [23,24], or to detect adulterations in egg-pasta [25].

Among the hundreds of volatiles compounds of pasta flavour, the multivariate statistical analysis allowed the selection of a limited number of aromatic substances useable as markers. Therefore, when food safety and nutritional intake are guaranteed, the flavour can become a discriminating factor to prefer a product rather than any other obtained with the same raw materials but in different process conditions. The identification and the quantification of these volatile compounds - to be assessed together with the conventional parameters (e.g. colour, cooking resistance, moisture content) - could represent a useful tool to promote and valorise pasta which has not undergone extreme thermal treatments during drying step, that includes those pasta productions that despite having been produced in industrial plants, yet respect Italian tradition of the LT-St process.

2. Materials and methods

2.1. Samples

All considered samples were short-shaped dried pasta. Pasta samples were all produced in Italy and sold from popular brands on the territory (leader brands, common brands, private labels). Samples were purchased from large-scale retail stores and from shops specializing in artisanal and regional food specialities. All samples were packed in a plastic box. The selection of samples was based on the claims reported on the label by the producer (e.g. "*slow drying at low temperature*"; "*slow-dried*", "*low-temperature drying*"), selling price (ranged from 1 to 5 €/Kg approximately), and some organoleptic features directly evaluable from the package (e.g. colour, surface appearance). The sample-set consisted of 52 pasta produced under on HT/VHT-St and LT-Lt drying process. Before analysis, samples were kept in their original packages at room temperature. Samples were fine ground for 30 s using an IKA A 10 laboratory mill to obtain a flour. The milling chamber was cooled with tap water to avoid heating during milling and thus loss of volatile compounds. A ground sample of 10.5 g was transferred into a 20 mL glass vial for autosampler with an aluminium-crimp top closure and blue-silicone/PTFE septum (Chromacol).

2.2. HS-SPME/GC-MS analysis

HS-SPME technique was carried out with a 50/30 Divinylbenzene/Carboxen/Polydimethylsiloxane (DVB/CAR/PDMS) Stable Flex fibre (2 cm) by Merch (KGaA, Darmstadt, Germany) using an autosampler device (Triplus Autosampler, Thermo Scientific). The vial with the

ground sample was automatically transferred in the autosampler oven preheated at 75 °C and maintained alike for 20 min. After the incubation period, the fibre was exposed in the headspace at the same temperature for 38 min. Following the extraction phase, the fibre was transferred automatically by the autosampler to the split/splitless injector of the gas chromatograph (Trace 1300 Gas Chromatograph, Thermo Scientific). The desorption step was performed in splitless mode (5 min) with a split flow of 50 mL/min and by setting the injector temperature at 260 °C. The chromatographic separation was performed with VF-WAXms capillary column (30 m × 0.25 mm ID, 0.25 mm) by Agilent Technologies (Santa Clara, CA, USA) with helium as carrier gas at a flow rate of 1 mL/min. The GC oven temperature was programmed at 40 °C for 5 min, ramped at 6 °C/min to 150 °C, then at 15 °C/min to 230 °C held for 3 min. The GC was interfaced with a mass spectrometer (ISQ 7000 Single Quadrupole, Thermo Fisher Scientific). The detection was performed under electron impact (EI) ionisation at 70 eV by operating in the full-scan acquisition mode in the 35–350 a.m.u. Range. The temperature of both the ion source and the transfer line was maintained at 250 °C. All pasta samples were analysed in duplicate following the random order required by the experimental design. The sample vials were interspaced with empty vials in order to avoid cross-contamination.

2.3. Experimental design

To optimize the extraction conditions, a strategy based on a rational design of experiments was followed. In particular, having identified three experimental factors as possibly critical to determine the extraction recovery and hypothesizing that there could be a non-linear relationship between the value of at least some of the factors and the response(s) to be optimized, a strategy based on selecting the experiments to be conducted by means of a face-centered central composite scheme, in order to build a response surface, was adopted. Central composite design (CCD) [26] are a family of experimental design obtained by the combination of a 2 k full factorial scheme (with experiments conducted at all the possible points resulting by taking the factors at their minimum and maximum levels) with a star design (where experiments are made under conditions in which all factors but one are kept constant at their mean level, while the remaining one assumes the values $+\alpha$ or $-\alpha$, plus a further experiment at the center point of the experimental domain). If the value of α, in coded coordinates, is set to 1, each factor is controlled at three levels and the CCD is said to be face-centered (FC). For a problem involving three factors, like the one addressed in the present study, a FC-CCD requires to conduct 15 experiments (Table 1). Such a number of experiments allows estimating reliably the values of the coefficients of the response surface:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 + b_{11}x_1^2 + b_{22}x_2^2 + b_{33}x_3^2 \quad (1)$$

which describes the relation between the controlled factors (and their binary interactions) and the response(s). The regression coefficients b_i can be considered as an estimate of the effect induced by the corresponding design term on the response. Accordingly, to evaluate whether a particular design term has a significant effect on the response or not, a *t*-test is used to verify whether the corresponding coefficient significantly different than zero. If replicated measurements are performed to estimate the repeatability of the response, this translates to calculating a critical value of the effect (at a specific confidence level). When, as in the present case, the whole design is replicated and a type-I error of 0.05 is chosen for the two-tailed *t*-test, the critical value for an effect to be considered significant is:

$$E_{crit} = t_{0.975,\,N}\frac{s_y}{\sqrt{N}} \quad (2)$$

where $t_{0.975,\,N}$ is the value of the t distribution of N degrees of freedom corresponding to a cumulative distribution function of 0.975, N is the number of points in the design and s_y is the pooled standard deviation of the response, estimated through the replication of the design points. Accordingly, if $|b_i| > E_{crit}$ the corresponding design term is assumed to affect significantly the response [27].

In the present study, as described in more detail in the Sub-section "*Optimization of extraction conditions through design of experiments*", both a targeted and an untargeted approach have been followed to define the responses to be optimized. In the first case, the peak areas of eight target analytes (hexanal, 2-pentylfuran, 1-hexanol, furfural, 1-octen-3-ol, 2-furanmethanol, maltol, nonanoic acid), important for pasta flavour, have been chosen as responses and further summarized into a desirability function [28]. In the second one, the TIC profiles collected across the experiments have been suitably pretreated and subjected to principal component analysis, and the corresponding scores along the first PC were used as response.

Table 1
Experiments of the FC-CCD for volatiles extraction in coded and real values.

Experiment nr.	Temperature		Time		Sample weight		Response (targeted) [a]		Response/10 [9] (untargeted) [b]	
	Coded	Real (°C)	Coded	Real (min)	Coded	Real (g)	Repl.1	Repl.2	Repl.1	Repl.2
1	−1	40	0	38	0	7.25	0.175	0.144	−3.99	−4.42
2	1	75	0	38	0	7.25	0.720	0.616	6.24	4.97
3	0	58	−1	15	0	7.25	0.261	0.221	−1.59	−2.32
4	0	58	1	60	0	7.25	0.427	0.369	3.01	2.44
5	0	58	0	38	−1	4	0.399	0.259	−0.73	−1.74
6	0	58	0	38	1	10.5	0.332	0.384	1.76	2.57
7	0	58	0	38	0	7.25	0.337	0.307	−0.70	0.27
8	−1	40	−1	15	−1	4	0.128	0.106	−5.36	−5.74
9	−1	40	−1	15	1	10.5	0.177	0.134	−4.06	−5.01
10	−1	40	1	60	−1	4	0.148	0.158	−4.68	−4.15
11	−1	40	1	60	1	10.5	0.160	0.151	−4.05	−3.87
12	1	75	−1	15	−1	4	0.366	0.318	0.35	−0.12
13	1	75	−1	15	1	10.5	0.518	0.389	3.54	1.77
14	1	75	1	60	−1	4	0.746	0.661	5.80	5.56
15	1	75	1	60	1	10.5	0.848	0.756	7.57	6.68

[a] Overall desirability calculated from the peak areas of hexanal, 2-pentylfuran, 1-hexanol, furfural, 1-octen-3-ol, 2-furanmethanol, maltol, nonanoic acid.
[b] Score on the first PC calculated on pretreated TIC chromatographic profiles.

2.4. Chemometric processing of chromatographic data

The possibility of using the HS-SPME/GC-MS profile for verifying whether a sample has been manufactured using as LT-Lt or a HT-St drying step and at the same time the identification of possible markers of the treatments relies on the construction and validation of a chemometric classification model [29,30]. In particular, given the highly multivariate nature of the chromatographic signals, in the present work the choice fell on partial least squares discriminant analysis (PLS-DA) [31,32]. Indeed, deriving from the partial least squares (PLS) regression algorithm [33], PLS-DA allows obtaining accurate and stable models also in cases, as the present one, when predictors are correlated, and the number of samples is lower than the number of measured variables. Briefly, in PLS-DA the categorical information about the samples is encoded in a dummy response variable which, when there are only two classes, like in the present study, can only take the values 1 (class 1) or 0 (class 2). Accordingly, the chromatographic profiles (collected in the matrix X) and the dummy response for the corresponding samples (collected in the column vector y) are used to build a regression model by means of the PLS algorithm:

$$y = \hat{y} + e = Xb + e \tag{3}$$

where b is the vector of regression coefficients, $\hat{y}$ collects the predicted responses for all the samples and the residuals (i.e., the differences between the actual and predicted values of the response) constitute the vector e. Since the predicted responses are not discrete-valued as their actual target, classification requires setting a proper threshold to the value of $\hat{y}$, so that if $\hat{y}$ is greater than the threshold, the sample is predicted as class 1, and otherwise it is assigned to class 2. In the present work, among the possible decision criteria, thresholding was operated by applying linear discriminant analysis on the values of the predicted response [34]. In particular, having assumed that prior probabilities for the two classes should be equal, since there was no reason for one treatment to be, a priori, more probable than the other, for each of the two categories, the posterior probability $p\left(g\hat{y}_i\right)$ was calculated based on the PLS predicted response of the training samples as:

$$p\left(g\hat{y}_i\right) = C_g e^{-\frac{\left(\hat{y}_i - \bar{y}_g\right)^2}{2s_g^2}} \quad g = 1,\ 2 \tag{4}$$

where $\bar{y}_g$ and s_g^2 are the mean and variance of the predicted responses for the training samples of the gth class, respectively, and C_g is a normalization factor. Accordingly, the classification threshold was calculated as the value of the PLS predicted response $\hat{y}_i$ corresponding to equal a posteriori probability for the two categories: $p\left(1\hat{y}_i\right) = p\left(2\hat{y}_i\right)$.

Once a predictive model is built, it needs to be validated, so that its reliability and generalizability to new samples can be evaluated [35]. In the present study, in order to take into account, the not very large number of samples available and at the same time to guarantee an unbiased validation of the results, a repeated double cross-validation (rDCV) procedure was used for this purpose [36]. Double cross-validation (DCV) consists of two nested cross-validation loops, the innermost one being used for model selection (e.g., choice of the optimal number of latent variables) and the outermost one for the evaluation of the predictive ability on external (i.e., not used for training) samples. The term "repeated" come from the fact that, in order for the results not to depend on a single data split, the procedure is repeated a sufficient number of times (here, 50), each time changing the way samples are divided into the different cancelation groups. To further rule out the possibility that good results may be obtained by chance correlation, the values of the figures of merit obtained by the rDCV procedure were tested against their corresponding null distributions, obtained non-parametrically by means of permutation tests [37].

Since chromatographic data may be affected by other sources of unwanted variance which may mask the information of interest (in the present case, the difference in volatiles' concentration between LT-Lt and HT-St dried pasta samples), prior to building and validating the classification model, signals need to be pretreated. Accordingly, alignment of the TIC chromatograms was first carried out using the iCoshift algorithm which, with respect to other warping procedures, has the advantage of allowing a finer matching of the profiles, by working interval-wise [38]. Then, normalization of the signals was achieved using probabilistic quotient normalization (PQN). PQN operates by identifying a reference profile which is usually the mean or the median signal, and then, for each sample, calculates the normalization factor as the median of the ratio between signal of the sample and that of reference [39].

3. Results and discussion

The results obtained by HS-SPME/GC-MS chromatographic profiles and processed using chemometric classification models showed it could be possible to use the flavour fingerprints of end products to recognize pasta which has been undergone thermal stress during its manufacturing process. On the market pasta packs reporting misleading information can be found claims such as *slow processing* or *artisanal product* are frequently used although they refer to secondary stages of the manufacturing process or their artisanal designation refers to the pasta factory only. Consequently, such label information does not automatically attest a top-quality pasta. On the other hand, pasta packs, produced in industrial plants that operate continuously on a large-scale (big brands or private label of large supermarket chains) but manufactured under LT-Lt drying treatments, can also be found. This procedure could then offer an objective tool to both commercially valorise the pasta produced with a LT-Lt drying technology and to help consumers make their conscious choice.

Before confirming data through the chemometric analysis, the qualitative assessment of chromatographic profiles, obtained comparing the volatiles compounds fingerprint of LT-Lt and HT/VHT drying process pasta, already showed differences in composition and amount of some characteristic substances of the two different drying technologies used for their production. As known, HT/VHT treatments promote the development of the Maillard reaction and, thus, several volatile molecules arising from either the early or the advanced stage of this reaction such as maltol, furfural, 2-furan-methanol, and benzaldehyde, are abundant in pasta samples that do not report information regarding low temperatures on the label, so, presumably, they had been dried rapidly at high or very high temperatures. Conversely, the LT drying process, that does not require sufficient temperatures to promote Maillard reaction, leads to the formation of molecules arising from the oxidation of polyunsaturated fatty acids, such as linear aldehydes and alcohol, volatile acids, 2-pentyl furan, largely detected in LT-Lt drying process samples. The abundance of the by-products of lipid oxidation compared to Maillard by-products could indicate that those samples were dried at $T < 68$ °C, because above this temperature, the enzymes (lipoxygenase, peroxidase, lyase hydroperoxide), that catalyse lipid oxidative processes, are denatured [40,41]. The qualitative results were obtained by comparing the experimental mass spectra with those stored in the US-NIST library database. According to the accepted criteria defined by the Metabolomics Standards Initiative [42], qualitative results based on spectral similarity with commercial mass spectral database, without use of reference standards, should be reported as *putative compounds identifications* (Level 2). Identification by comparison with mass spectra was considered satisfactory only for compounds with obtained spectra reverse match (R match) higher than 850. As spectral evidence to confirm the compounds identifications, the Kovats retention indices (RI) were determined using a homologous series of aliphatic hydrocarbons

(C8–C25) (Sigma, Aldrich, Milan, Italy) as external references and matched with those available in literature [43,44].

3.1. Optimization of extraction conditions through design of experiments

In the first phase of the work, starting from a method previously developed by some of the co-authors [18], experimental design was used to try to improve the extraction yields of the analytes of interest (but also of the whole panel of volatiles, in general), by investigating the effect of the sample weight, and of the extraction temperature and time on the recovery.

Accordingly, a three-factor face centered central composite design (FC-CCD) was constructed, identifying the limits of the experimental domain as the values 40 and 75 °C for sampling temperature, 15 and 60 min for sampling time and 4 and 10.5 g for sample weight. Starting from these conditions, the design matrix reported in Table 1 and comprising 15 experiments was then generated. To assess the precision of the measurements, so to be able to evaluate the significance of the effects of each design term, it was decided to perform all the required experiments in duplicate; in particular, to avoid the possibility of highlighting spurious correlations or masking effects, the experiments were carried out in random order, taking care that the randomization also extended to the replicated measures.

In order to evaluate the recoveries, the 30 (15 × 2) extracts prepared according to the points of the experimental design were analysed by GC-MS as described in the Section "*HS-SPME/GC-MS analysis*". The corresponding chromatographic data were imported into Matlab, aligned using the iCoshift algorithm, and subsequently processed. In particular, to define the responses to be optimized based on the results of the experimental design two different approaches were followed: a targeted and an untargeted one.

The former approach relied on the quantification of eight substances (hexanal, 2-pentylfuran, 1-hexanol, furfural, 1-octen-3-ol, 2-furanmethanol, maltol, nonanoic acid) which are reported in the literature as important for the constitution of the aroma of the pasta. For each of these substances, the areas of the corresponding peaks were recorded and, in order to take into account the recovery of all eight chosen markers and their different concentrations within the samples, starting from those areas a desirability function [28] was built as follows. Since, for each compound, the aim was to maximize the recovery, i.e., to have the highest peak area, individual desirability values were calculated for each compound, by means of the following function:

$$d_k\left(A_i^{(k)}\right) = 0.8 \times \frac{A_i^{(k)} - L_k}{U_k - L_k} + 0.1 \quad (5)$$

where $d_k\left(A_i^{(k)}\right)$ is the value of the desirability function for the kth response in the ith experiment, $A_i^{(k)}$ is the area of the kth analyte in the ith experiment, and L_k and U_k are the minimum and maximum value of the kth response across all the experiments. Accordingly, through the normalization in equation (5), the peak areas for each analyte were transformed to an individual desirability value, which is normalized so to vary between 0.1 and 0.9. Then, for each experiment, the overall desirability function was calculated as the geometric mean of the individual desirabilities:

$$D_i = \prod_{k=1}^{N_{analytes}} \left[d_k\left(A_i^{(k)}\right)\right]^{\frac{1}{N_{analytes}}} \quad (6)$$

The results are reported in Table 1 and graphically summarized in Fig. 1.

To facilitate the interpretation of the results, the experiments reported in Fig. 1 have been organized in the order described in Table 1, the outcomes of replicated design units being plotted consecutively and labelled as a and b. By looking at Fig. 1, it is then possible to observe, at first, how all the extractions have good repeatability in recoveries, since the desirability values corresponding to pairs of replicated experiments are very consistent with one another. Furthermore, it is clear that, among those examined, three experimental conditions result in overall higher recoveries: these conditions correspond to the experiments indicated in Table 1 as 15 (75 °C–60 min – 10.5 g), 14 (75 °C–60 min – 4 g) and 2 (75 °C–38 min – 7.25 g).

A first inspection of these results would suggest the need to conduct the experiments at high temperature and for a medium-long time, but it would also indicate, on the other hand, how the quantity of sample used may not be a significant factor for obtaining high recoveries.

As said, the outcomes of the designed experiments were analysed also through a second, untargeted approach, taking into consideration

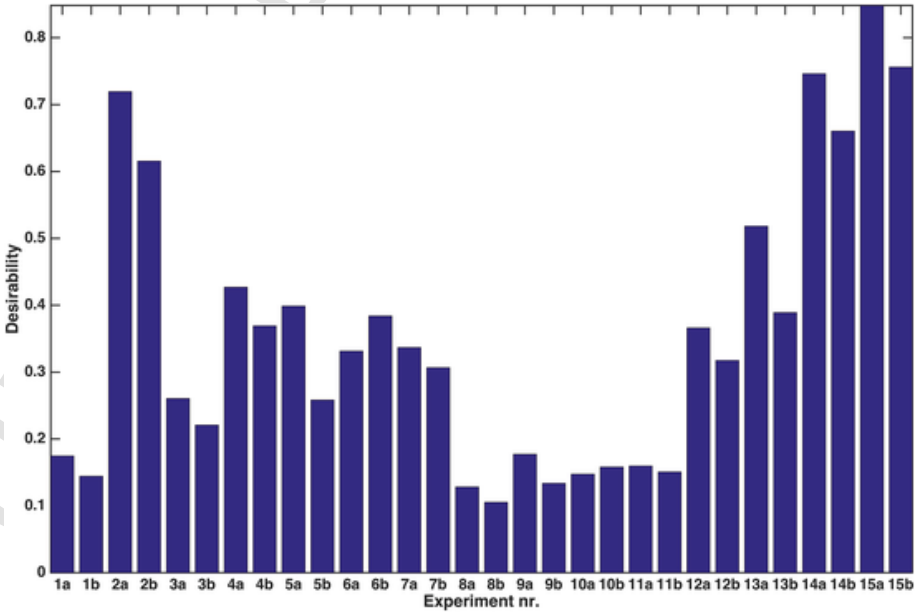
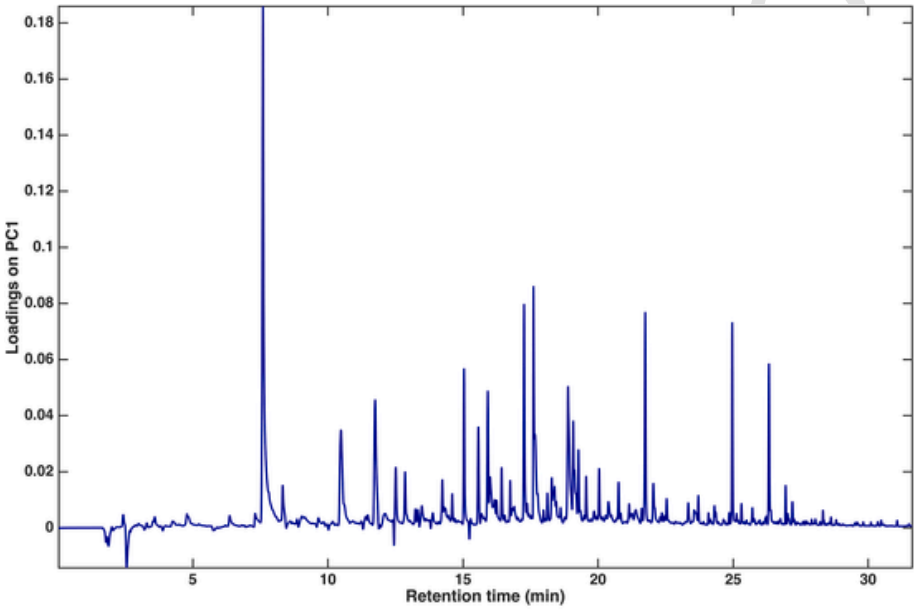


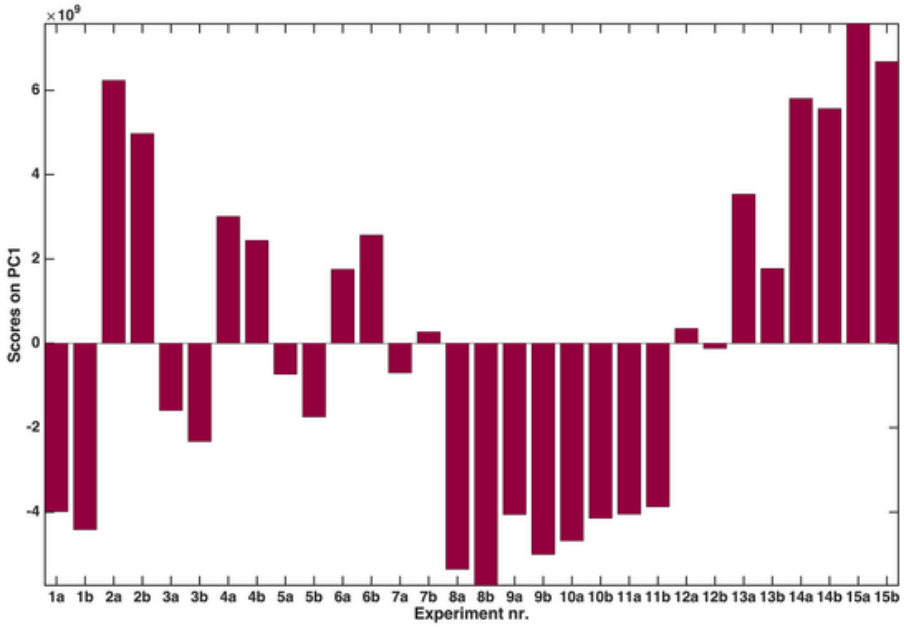
Fig. 1. Optimization of the extraction conditions based on a targeted approach: values of the desirability functions for the two replicates (labelled a and b) of the fifteen points of the experimental design.

the entire chromatograms and not just the areas of specific peaks of interest. To do so, the aligned chromatograms, after further mean centering, were processed by principal component analysis. In particular, inspection of the loadings of the first principal component (Fig. 2a), which accounts for more than 70% of the original variance, evidenced how, with the exception of a small region at the beginning of chromatogram, very close to the dead volume, the contributions of all the peaks are always positive: this means that the scores along the corresponding component (Fig. 2b) can be considered as an index of the overall recovery of eluted volatile substances. Indeed, based on what reported in Fig. 2, it is possible to say that the higher the intensity of the chromatographic peaks, the higher the score value on PC1. By looking at the values of the scores along PC1 for the different experimental conditions (Fig. 2b), it is possible note how the observed trend is consistent with the one already discussed for the desirability function built considering only the 8 markers identified for the targeted analysis (Fig. 1).

Based on these considerations, since the scores on the first principal component calculated on the whole chromatographic profile could be considered as an index of the extraction recovery of all volatile substances of interest (and not only of the 8 selected markers), it was decided to use this variable as response to build the quantitative model to



a



b

Fig. 2. Optimization of the extraction conditions based on an untargeted approach. Results of PCA modelling of the TIC chromatograms measured on the 30 (15 × 2) experiments selected according to the Face Centered-Central Composite Design. Loadings (a) and scores (b) of the first principal component.

be used for the optimization of the experimental conditions. By indicating as y the response to be optimized (i.e., the score along PC1), the equation of the response surface to be fit based on the results of the experimental design is:

$$y = b_0 + b_T T + b_t t + b_w w + b_{T^2} T^2 + b_{t^2} t^2 + b_{w^2} w^2 + b_{Tt} Tt + b_{Tw} Tw + b_{wt} wt \quad (7)$$

where T, t and w are the values of temperature, time and sample weight, respectively. Having replicated each point of the experimental design, it was possible not only to estimate the values of the regression coefficients b_i in equation (7), but also their statistical significance. The results obtained are summarized in Fig. 3.

By looking at Fig. 3, it is possible to see how the effect of temperature is predominant and positive; similarly, also time and the interaction between temperature and time have a positive and significant effect, although lower. There results confirm the qualitative considerations already discussed on the basis of Figs. 1 and 2. As for the sample weight, it can be deduced that it has a minimal effect, only additive, since its interactions and its quadratic term do not give statistically significant contributions.

Accordingly, considering only the statistically significant contributions, the mathematical model describing the response surface and reported in equation (7) boils down to:

$$y = b_0 + b_T T + b_t t + b_w w + b_{T^2} T^2 + b_{t^2} t^2 + b_{Tt} Tt \quad (8)$$

In order to choose the optimal conditions, taking advantage of the fact that the sample weight is not present in quadratic or interaction terms, but only through the linear term, it can be evaluated separately from the contribution of the other two factors. In fact, considering that the mass gives only a linear contribution with a positive coefficient, it can be concluded that, in order to have a response as high as possible, it must be fixed at its maximum level.

At this point, it is possible to proceed by investigating the marginal response surface, obtained from equation (8) by fixing the sample weight at its optimal level (10.5 g) and considering only the dependence of y on the remaining two factors, temperature and time. The corresponding response function is graphically shown in Fig. 4, both as surface and in the form of a contour plot.

By looking at Fig. 4, it is evident how the optimal conditions, i.e., the ones leading to the highest recovery of volatile substances, as summarized by the score on PC1 which was chosen as response, correspond to the highest values of time and temperature (i.e., considering the investigated experimental domain, to 75 °C for 60 min). However, in the light of trying to reduce time and costs of the analysis, a further investigation of the response surface highlighted how very good recoveries could still be obtained by keeping the temperature high, but lowering the extraction time to 38 min, so that these were chosen as the final experimental conditions. Accordingly, for the successive analysis of the pasta samples, extraction was performed on 10.5 g of sample at a temperature of 75 °C for 38 min.

3.2. Discrimination of pasta samples according to the processing conditions

The TIC chromatograms of the 52 analysed pasta samples (made up of 9301 data points each) were exported into Matlab, aligned by means of the iCoshift algorithm and normalized by means of probabilistic quotient normalization (PQN) using the median profile over the whole data set as reference chromatogram, as described in "*Chemometric processing of chromatographic data*" Section.

A PLS-DA classification model was then built and validated on the pretreated data, in order to discriminate between the two types of pasta (pasta produced under on HT/VHT-St and LT-Lt drying process). In particular, considering the number of available samples, validation has been performed through the combination of repeated double cross-validation (rDCV) and permutation test. More specifically, the rDCV procedure was implemented by dividing the samples in both loops into 10 cancellation groups and replicating the procedure for a total of 50 times. Permutation tests (with 1000 randomizations) were then used to obtain the distribution of the classification figures of merit under the null hypothesis in order to rule out the possibility that the results obtained could be due just to chance correlations. In a permutation test, the class labels of the samples are randomly assigned to the different individuals and do not match anymore the true grouping. Classification models are then built based on the permuted class labels, and the results are collected to represent the distribution of values one should expect in cases where there is no difference between the categories [37].

The results of PLS-DA modelling were very good in terms of predictive ability; here it should be stressed that the use of a repeated double

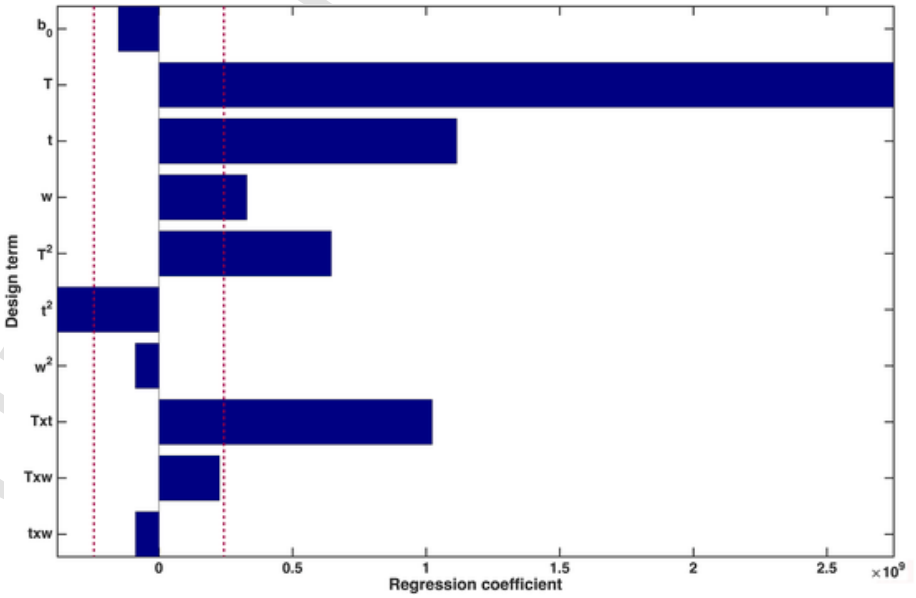


Fig. 3. Optimization of the extraction conditions based on an untargeted approach. Regression coefficients defining the response surface. The vertical dashed red lines indicate the threshold for statistical significance (p = 0.05). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

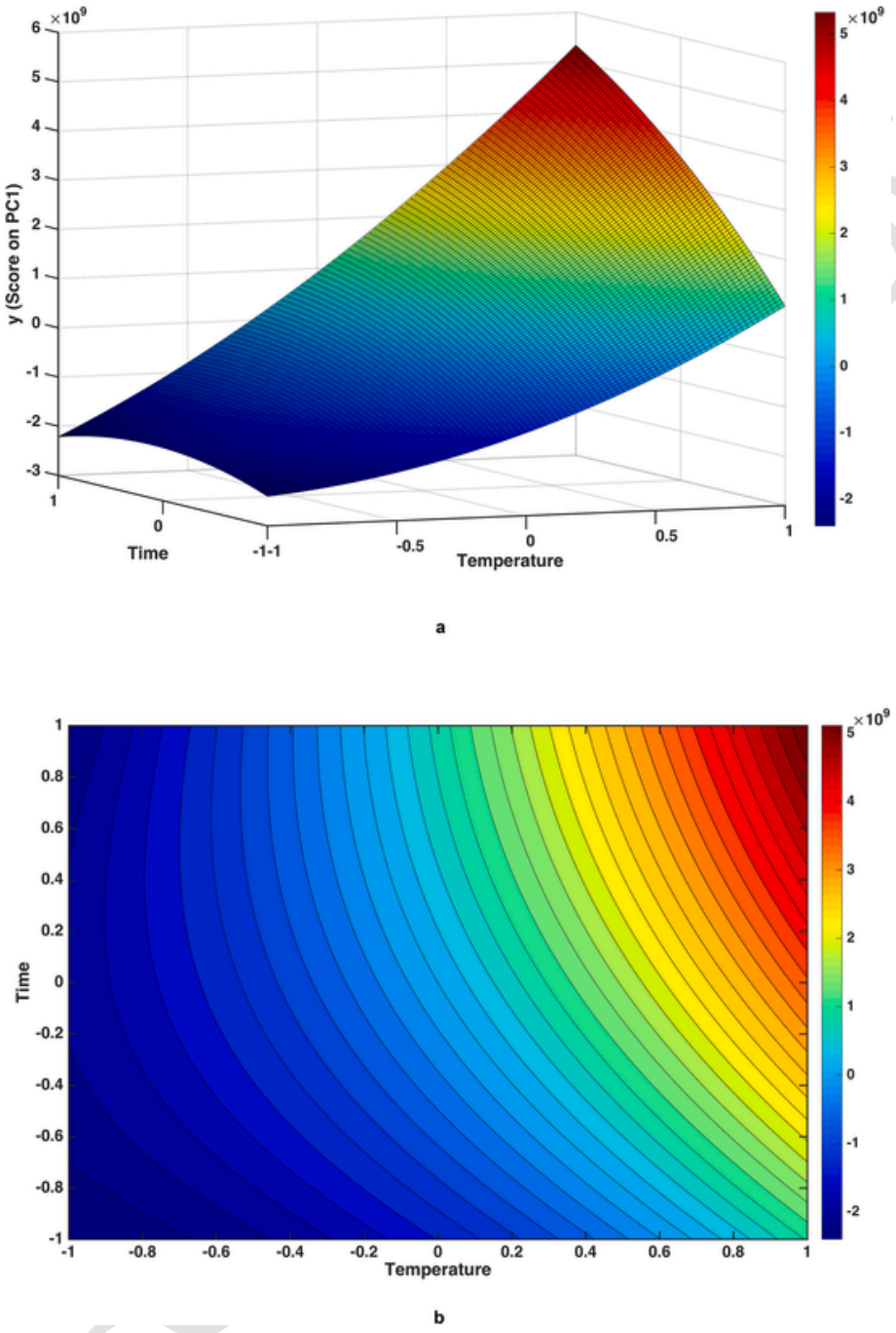


Fig. 4. Optimization of the extraction conditions based on an untargeted approach. Marginal response surfaces illustrating the dependence of the modelled response (scores on PC1) as a function of headspace extraction time and temperature. (a) 3D response surface; (b) contour plot.

cross-validation strategy allows to obtain not only a single estimate of the correct classification rates of the individual classes or the whole data set, but also their confidence intervals.

In particular, it was found that the PLS-DA method allows, on average, correctly predicting the category of more than 80% of the outer loop samples, i.e., when they are not employed to build the model for their prediction. Indeed, due to the nature of the rDCV procedure the figures of merit estimated on the outer loop samples are the ones which estimate less unbiasedly the performances which can be expected on new, unknown samples. In detail, when considering the outer loop, a 74.0 ± 8.2% sensitivity and an 86.5 ± 3.9% specificity were obtained for the class LT-Lt, while, due to the symmetry of the two-class discriminant problem, the values were exchanged in the case of the HT/VHT-St category (86.5 ± 3.9% sensitivity and 74.0 ± 8.2% specificity), leading to an overall correct classification rate (accuracy) of 80.3 ± 4.7%. Moreover, the fact that these results were comparable to the values of the same figures of merit observed in the inner loop (the one used for model selection, where the sensitivities were 74 ± 10% and 89.4 ± 3.0% and, correspondingly, the specificities 89.4 ± 3.0% and 74 ± 10% for LT-Lt and HT/VHT-St, respectively, with an accuracy of 81.5 ± 5.6%) indicate the absence of overfitting. By inspecting the reported results, it is also possible to observe how the sensitivity for HT/VHT-St dried pasta presents a smaller standard deviation, indicating that, probably, the corresponding class is more homogeneous than the alternative one (LT-Lt).

Apart from the individual sensitivity and specificity, the goodness of the classification model can be evaluated and summarized by three other figures of merit: the number of incorrect classifications (Number of Misclassifications, NMC), the area under the ROC curve (AUROC) and the discriminant Q^2 (DQ2) [37]. The number of misclassification is simply equal to the number of samples which are predicted as belonging to the wrong category (HT/VHT-St classified as LT-Lt and vice versa). The receiver operator characteristics (ROC) curve is a way to graphically summarize the classification performances of a method as a function of the values of the discriminant threshold: the curve plots the sensitivity of a class versus 1-specificity (for a two-class problem, it is enough to consider only one of the two categories involved due to symmetry: the sensitivity of a class is equal to the specificity of the other and vice versa). The ROC curve doesn't consider only the performances at the optimal value of the classification threshold but shows how sensitivity and specificity would vary by changing the decision boundary, thus providing more information about the ability of the classifier to differentiate the two categories. As a figure of merit, the area under the ROC curve (AUROC) is often used: it takes the value of 1 for a perfect separation between the classes, while no separation corresponds to a value close to 0.5. Lastly, discriminant Q^2 (DQ2) was introduced as a modification of the coefficient of determination to evaluate the quality of regression-based classification models [45]. DQ2 is defined as the usual Q2 but without considering, in the calculation of the residual sum of squares, the errors associated to predictions of values beyond the class labels: if for class 1 $\hat{y}_i>1$ or for class 2 $\hat{y}_i<0$, these errors do not contribute to the residuals.

Concerning their values, while NMC should be as low as possible (and, ideally, equal to zero), the other two should be as high as possible (and, ideally, equal to 1). The values of these three figures of merit are represented in Fig. S1 and compared with their distributions under null hypotheses, which were non parametrically estimated by means of permutation tests with 1000 repetitions [37].

As explained above, the low NMC value (10.3), like the high value of AUROC (0.860), are indicators of the goodness of the model; moreover, by comparing the values obtained on the data with the corresponding distributions under the null hypothesis estimated by means of the permutation test, it is evident how both figures of merit are statistically significant (with an estimated p-value <0.001). On the other hand, inspection of the value of DQ2 (−1.26) and comparison of this value with its distribution under the null hypothesis provide less strong support to the observed discrimination but this is not unexpected, as there is debate in the literature about the suitability of this figure of merit in assessing the goodness of a classification model.

In order to identify which chromatographic regions contribute the most to the classification model and, accordingly which compounds allow discriminating between the two types of pasta, VIP (variable importance in the projection) indices were calculated for each variable. Indeed, the normalization adopted in the calculation of these indices is such that only the variables with VIP higher than 1 are assumed to contribute significantly to the classification model. Moreover, having identified the significant X variables (chromatographic regions), inspection of the sign of the PLS regression coefficients, which indicate the how the dummy Y variable coding for class belonging varies as a function of each predictor in the X matrix, it is possible to obtain information about how the signals from the markers vary between the two types of pasta (Fig. 5).

In particular, in Fig. 5 the chromatographic regions corresponding to VIP indices greater than 1 are highlighted by a bold trait, whereas the different colour is associated to the value of the regression coefficient: the red regions are those in which the LT-Lt dried pasta have a more intense signal, while the opposite occurs for the blue ones (the signals of HT/VHT-St products are more intense).

Among the blue peaks, it is possible to clearly identify those related to 2-furanmethanol and maltol: this is consistent with the characteristic of the HT/VHT-St pasta samples, for which the use of high temperatures implies the formation of these Maillard reaction by-products.

On the other hand, among the red variables it is possible to identify the presence of 1-hexanol, as expected, together with many other peaks ascribable to alcohols and carboxylic acids, produced by the oxidation of lipids, which is possible at the operating conditions of LT-Lt products.

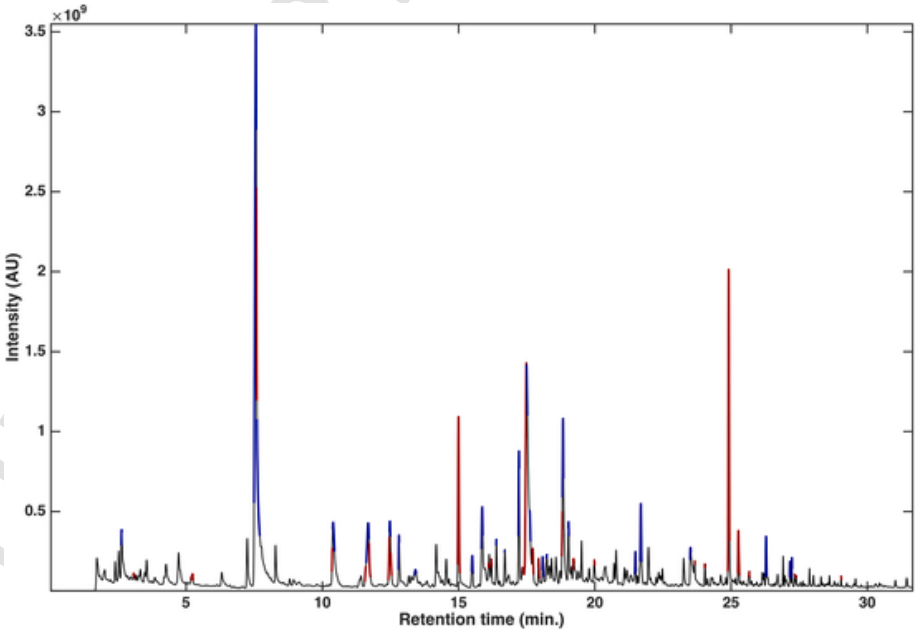


Fig. 5. Identification of potential marker based on the results of PLS-DA modelling. The chromatographic regions found to be significantly contributing to the model are highlighted by a bold trait over the average experimental profile of the analysed samples. Red and blue colours indicate, respectively, variables being on average more intense for LT-Lt or HT-St, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

4. Conclusions

The goal of the present study was to analyse durum wheat pasta samples by HS-SPME/GC-MS in order to investigate the chemical information provided by their flavour profiles. As known, in fact, the different thermal conditions used in drying treatments HT/VHT-St and LT-Lt have a direct impact on qualitative and quantitative composition of the end-product and, consequently, they could influence the nutritional and organoleptic pasta qualities. Samples set consisted in 52 short-shaped dried pasta commercialized by most common Italian brands. The selection of samples was based on the information reported on the label by the producer, i.e. whether the product was manufactured under HT/VHT-St or LT-Lt drying process. A multiresidue analytical method to investigate the volatile compounds in pasta samples was developed and the HS-SPME procedure was optimized using a FC-CCD. The chromatographic profiles were then used to classify samples according to the drying process their underwent. PLS-DA analysis (validated by means of repeated double cross-validation and permutation tests) allowed, on average, correctly predicting more than 80% of the samples, confirming that the drying step may have a significant impact on the pasta flavour. In addition, the presence of several volatile compounds arising from either the early or the advanced stage of the Maillard reaction (maltol, furfural, 2-furan-methanol, benzaldehyde) in pasta samples that do not report claim regarding low temperatures on the label, could confirm that those samples had been dried rapidly at high or very high temperatures. On the other hand, the abundance of alcohols and carboxylic acids (1-hexanol, 1-octen-3-ol, nonanoic acid) ascribable to the lipid oxidative processes in pasta produced under drying LT-Lt as stated on the label could hence confirm that the samples have not undergone to thermal stress. It was also observed that the pasta samples dried at high or very high temperature show a general flattening of the flavour fingerprint. Therefore, the HT-VHT/St dried pasta could have an organoleptic quality lower compared to those produced at low temperature.

In conclusion, the outcome indicates that targeted analysis of flavour profiles could be used as a tool to provide added-value to those pasta productions that, despite having been produced in industrial plants, respect the artisanal tradition of the LT-St process.

Author contribution

Vanessa Giannetti: Writing – original draft, Investigation, Methodology, Formal analysis, Project administration, Resources, Writing – review & editing. Maurizio Boccacci Mariani: Writing – original draft, Investigation, Methodology, Formal analysis, Project administration, Resources, Writing – review & editing. Federico Marini: Writing – original draft, Software, Formal analysis, Data curation, Writing – review & editing. Alessandra Biancolillo: Writing – original draft, Formal analysis, Data curation, Writing – review & editing.

Declaration of competing interest

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

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.talanta.2020.121955.

References

[1] Law No, 580 of 4 July 1967. Disciplina per la lavorazione e commercio dei cereali, degli sfarinati, del pane e delle paste alimentari. OJ No. 189 of 29/07/67. Amended by DPR (Decree of the President of the Italian Republic) No. 187 of 9 February 2001. Regolamento per la revisione della normativa sulla produzione e commercializzazione di sfarinati e paste alimentari. OJ No. vol. 117, 22/05/2001.
[2] M. Anese, C. Nicoli, R. Massini, C.R. Lerici, Effects of drying processing on the Maillard reaction in pasta, Food Res. Int. 32 (1999) 193–199.
[3] P. Resmini, L. Pellegrino, Occurrence of protein-bound lysylpyrrolaldehyde in dried pasta, Cereal Chem. 71 (1994) 254–262.
[4] A. Sensidoni, D. Peressini, C.M. Pollini, M. Munari, Effect of mechanical stress and reducing sugar content of semolina on pasta damage during drying, Italian Food and Beverage Technol. 7 (1996) 9–14.
[5] B. Aktan, K. Khan, Effect of high-temperature drying of pasta on quality parameters and on solubility, gel electrophoresis, and reversed-phase high-performance liquid chromatography of protein components, Cereal Chem. 69 (1992) 288–295.
[6] G. Pavan, High temperature drying improves pasta quality, Food Eng. Int. 5 (1980) 37–39.
[7] R. Acquistucci, Influence of Maillard Reaction on protein modification and colour development in pasta. Comparison of different drying conditions, LWT - Food Sci. Technol. (Lebensmittel-Wissenschaft -Technol.) 33 (2000) 48–52.
[8] I. De Noni, M.A. Pagani, Cooking properties and heat damage of dried pasta as influenced by raw material characteristics and processing conditions, Crit. Rev. Food Sci. Nutr. 50 (2010) 465–472.
[9] M. Petitot, C. Barron, C. Larré, M.H. Morel, V. Micard, Modification of pasta structure induced by high drying temperatures. Effects on the in vitro digestibility of protein and starch fractions and the potential allergenicity of protein hydrolysates, Food Chem. 116 (2009) 401–412.
[10] M. Stuknytė, S. Cattaneo, M.A. Pagani, A.M.V. Micard, J. Hogenboom, I. De Noni, Spaghetti from durum wheat: effect of drying conditions on heat damage, ultrastructure and in vitro digestibility, Food Chem. 149 (2014) 40–46.
[11] R. Beleggia, C. Platani, R. Papa, A. Di Chio, E. Barros, C. Mashaba, J. Wirth, A. Fammartino, C. Sautter, S. Conner, J. Rauscher, D. Stewart, L. Cattivelli, Metabolomics and food processing: from semolina to pasta, J. Agric. Food Chem. 59 (17) (2011) 9366–9377.
[12] H.F. Erbersdobler, V. Somoza, Review. Forty years of furosine – forty years of using Maillard reaction products as indicators of the nutritional quality of foods, Mol. Nutr. Food Res. 51 (2007) 423–430.
[13] V. Giannetti, M. Boccacci Mariani, P. Mannino, Furosine as a pasta quality marker: evaluation by an innovative and fast chromatographic approach, J. Food Sci. 78 (2013) 993–999.
[14] E. Marconi, M. Carcea, M. Schiavone, R. Cubadda, Spelt (Triticum spelta L.) Pasta Quality: combined effect of flour properties and drying conditions, Cereal Chem. 79 (2002) 634–639.
[15] C. Perez-Locas, V.A. Yaylayan, The Maillard reaction and food quality deterioration, in: L. Skibsted, J. Risbo, M. Andersen (Eds.), Chemical Deterioration and Physical Instability of Food and Beverages. Part 1 Understanding and Measuring Chemical Deterioration of Foods and Beverages, Woodhead Publishing Series in Food Science, Technology and Nutrition, 2010, pp. 70–94.
[16] A. Sayaslan, O.K. Chung, P.A. Seib, L.M. Seitz, Volatile compounds in five starches, Cereal Chem. 77 (2000) 248–253.
[17] N. Sozer, A.C. Dalgıç, A. Kaya, Thermal, textural and cooking properties of spaghetti enriched with resistant starch, J. Food Eng. 81 (2007) 476–484.
[18] M. Boccacci Mariani, V. Giannetti, E. Testani, HS-SPME/GC-MS method to characterise the flavour of Italian pasta: potential application to assess the quality of the products, Food Anal. Methods 7 (2014) 64–72.
[19] V. Giannetti, M. Boccacci Mariani, P. Mannino, E. Testani, Furosine and flavour compounds in durum wheat pasta produced under different manufacturing conditions: multivariate chemometric characterization, LWT - Food Sci. Technol. (Lebensmittel-Wissenschaft -Technol.) 56 (2014) 15–20.
[20] A. Pasqualone, V.M. Paradiso, C. Summo, F. Caponio, T. Gomes, Influence of drying conditions on volatile compounds of pasta, Food Bioprocess Technol. 7 (2014) 719–731.
[21] A. Baiano, P. Siciliano, A.M. Taurino, D.S. Presicce, G. Gambacorta, C. Lamacchia, S. Pati, E. La Notte, Electronic nose and GC-MS to investigate the volatile component of an Italian traditional pasta, Cereal Foods World 52 (2007) 138–143.
[22] R. Beleggia, C. Platani, G. Spano, M. Monteleone, L. Cattivelli, Metabolic profiling and analysis of volatile composition of durum wheat semolina and pasta, J. Cereal. Sci. 49 (2009) 301–309.
[23] P. Firmani, G. La Piscopia, R. Bucci, F. Marini, A. Biancolillo, P.G.I. Authentication of, Gragnano pasta by near infrared (NIR) spectroscopy and chemometrics, Microchem. J. 152 (2020) 104339.
[24] V. Giannetti, M. Boccacci Mariani, P. Mannino, Characterization of the authenticity of Pasta di Gragnano Protected Geographical Indication through flavor component analysis by gas chromatography-mass spectrometry and chemometric tools, J. AOAC Int. 99 (2016) 1279–1286.
[25] A. Biancolillo, A. Santoro, P. Firmani, F. Marini, Identification and quantification of turmeric adulteration in egg-pasta by near infrared spectroscopy and chemometrics, Appl. Sci. 10 (2020) 2647.
[26] G.E.P. Box, K.G. Wilson, On the experimental attainment of optimum conditions, J. Roy. Stat. Soc. B 13 (1951) 1–45.
[27] D.C. Montgomery, Design and Analysis of Experiments, eighth ed., John Wiley and Sons, New York, 2013.
[28] G. Derringer, R. Suich, Simultaneous optimization of several response variables, J. Qual. Technol. 12 (1980) 214–219.
[29] A. Biancolillo, F. Marini, Chemometrics applied to plant spectral analysis, in: J. Lopes, C. Sousa (Eds.), Vibrational Spectroscopy for Plant Varieties and Cultivars Characterization, Comprehensive Analytical Chemistry, 80, Elsevier, Amsterdam, 2018, pp. 69–104 (chapter 4).
[30] M. Cocchi, A. Biancolillo, F. Marini, Chemometric methods for classification and feature selection, in: J. Jaumot, C. Bedia, R. Tauler (Eds.), Data Analysis for Omic Sciences: Methods and Applications, Comprehensive Analytical Chemistry, 82, Elsevier, Amsterdam, 2018, pp. 265–299.

[31] M. Sjöström, S. Wold, B. Söderström, PLS discriminant plots, in: E.S. Gelsema, L.N. Kanal (Eds.), Pattern Recognition in Practice, Elsevier, Amsterdam, 1986, pp. 461–470.
[32] L. Ståhle, S. Wold, Partial least squares analysis with cross-validation for the two class problem: a Monte Carlo study, J. Chemom. 1 (1987) 185–196.
[33] S. Wold, H. Martens, H. Wold, The multivariate calibration problem in chemistry solved by the PLS method, in: A. Ruhe, B. Kagstrom (Eds.), Proceedings of the Conference on Matrix Pencils, Lecture Notes in Mathematics, 1983, pp. 286–293.
[34] U.G. Indahl, H. Martens, T. Næs, From dummy regression to prior probabilities in PLS-DA, J. Chemom. 21 (2007) 529–536.
[35] F. Westad, F. Marini, Validation of chemometrics models - a tutorial, Anal. Chim. Acta 893 (2015) 14–24.
[36] P. Filzmoser, B. Liebmann, K. Varmuza, Repeated double cross validation, J. Chem. 23 (2009) 160–171.
[37] J.A. Westerhuis, H.C.J. Hoefsloot, S. Smit, D.J. Vis, A.K. Smilde, E.J.J. van Velzen, J.P.M. van Duijnhoven, F.A. van Dorsten, Assessment of PLSDA cross validation, Metabolomics 4 (2008) 81–89.
[38] G. Tomasi, F. Savorani, S.B. Engelsen, Icoshift: an effective tool for the alignment of chromatographic data, J. Chromatogr. A 1218 (2011) 7832–7840.
[39] F. Dieterle, A. Ross, G. Schlotterbeck, H. Senn, Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics, Anal. Chem. 7 (2006) 4281–4290.
[40] U. Schweiggert, A. Schieber, R. Carle, Inactivation of peroxidase, polyphenoloxidase, and lipoxygenase in paprika and chili powder after immediate thermal treatment of the plant material, Innovat. Food Sci. Emerg. Technol. 6 (2005) 403–411.
[41] M.K. Sucan, D.K. Weerasinghe, Process and reaction flavors: an overview, ACS Symposium Series, 905, 2005, pp. 1–23 (Chapter 1).
[42] L.W. Sumner, A. Amberg, D. Barrett, M.H. Beale, R. Beger, C.A. Daykin, T.W.M. Fan, O. Fiehn, R. Goodacre, J.L. Griffin, T. Hankemeier, N. Hardy, J. Harnly, R. Higashi, J. Kopka, A.N. Lane, J.C. Lindon, P. Marriott, A.W. Nicholls, M.D. Reily, J.J. Thaden, M.R. Viant, Proposed minimum reporting standards for chemical analysis, Metabolomics 3 (2006) 211–221.
[43] T. Acree H. Arn Flavornet. Gas chromatography-olfatometry (GCO) of natural productshttp://www.flavornet.orgs 21 November 2020
[44] V.I. Babushok, P.J. Linstrom, I.G. Zenkevichb, Retention Indices for frequently reported compounds of plant essential oils, J. Phys. Chem. Ref. Data 40 (2011) 043101.
[45] J.A. Westerhuis, E.J.J. van Velzen, H.C.J. Hoefsloot, A.K. Smilde, Discriminant Q2 (DQ2) for improved discrimination in PLSDA models, Metabolomics 4 (2008) 293–296.