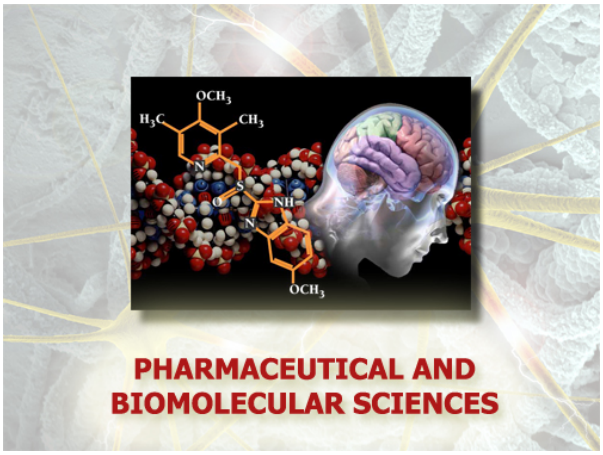
Università degli Studi di Torino

Scuola di Dottorato in
Scienze della Natura e Tecnologie Innovative

Dottorato in
Scienze Farmaceutiche e Biomolecolari
(XXX ciclo)



MULTIDISCIPLINARY APPROACHES IN FOODOMIC STUDIES: analytical and chemometric relationships between sensory evaluation and chemical composition of different coffees

Candidato: Davide Bressanello

Tutor: Dr. Erica Liberto, PhD

Università degli Studi di Torino

Dottorato in
Scienze Farmaceutiche e Biomolecolari

Tesi svolta presso il
Dipartimento di Scienza e Tecnologia del Farmaco

XXX

Multidisciplinary approaches in foodomic studies: analytical and chemometric relationships between sensory evaluation and chemical composition of different coffees

CANDIDATO: Davide Bressanello
TUTOR: Dr. Erica Liberto, PhD
COORDINATORE SCUOLA DOTTORATO: Prof. Dr. Gianmario Martra

ANNI ACCADEMICI: 2014/15
2015/16
2016/17

SETTORE SCIENTIFICO-DISCIPLINARE DI AFFERENZA: CHIM10

Acknowledgments

Davide Bressanello is indebted with “Luigi Lavazza s.p.a” (Torino, Italy) R&D Division for the financial support to his PhD thesis.
Erica Liberto, Carlo Bicchi and Davide Bressanello are grateful to Dr. Gloria Pellegrino and Dr. Manuela Rosanna Ruosi for helpful discussion and advice.

Davide Bressanello is grateful to Prof. Dr. Peter Schieberle for the guest period at the DFA, the helpful discussion and advice.

SUMMARY

ABSTRACT

The detailed investigation of food chemical composition has become a crucial task because of the remarkable increase of consumers' demand not only about safety, but also about traceability and sensory pleasant products.

The needs to satisfy this extended concept of quality encouraged the migration from conventional analytical approaches to new and more comprehensive strategies.

Among the "-omics" disciplines "*Foodomics*"[1,2] is the analytical approach that, in food chemistry, aims at an even more global characterization of the food to define its chemical, physical, sensory and nutritional properties in order to link them to its biological impact.

The particular branch of foodomics focused on linking the chemical composition to the sensory properties of a food assumes particular relevance in the case of comfort foods (such as coffee), whose choice and consumption is more driven by the pleasure given by its intake than by its nutritional properties.

This doctorate thesis aims to investigate the relationship between the chemical composition of coffee flavour and its sensory impact through an "-omic" approach.

In addition, a considerable importance has been given to the exploitation of this relationship to develop a tool suitable to be integrated with the routine coffee sensory evaluation.

In the first part, three different HCC sampling approaches combined with gas-chromatography-mass spectrometry have been used to simulate the three different steps of the official SCAA sensory evaluation protocol. Collected chemical data have been compared to investigate their consistency and if, as actually happens during the sensory analysis, they are all necessary to fully characterize coffee samples with macroscopic sensory differences.

HS-SPME-GC-MS of the roasted coffee powder demonstrated to be a highly satisfactory platform both in terms of characterizing power and possibility of automation.

Then a pool of samples, suitable to define the evolution of seven different sensory notes, have been analysed with the aforementioned technique and the collected chemical fingerprints used to develop a data elaboration workflow with the final aim to define predictive models able to attribute to the six sensory notes used to describe a given sample, a sensory score comparable to those measured by the panel.

In this part of the project six different sensory notes have been modelled with encouraging results and acceptable model performance parameters.

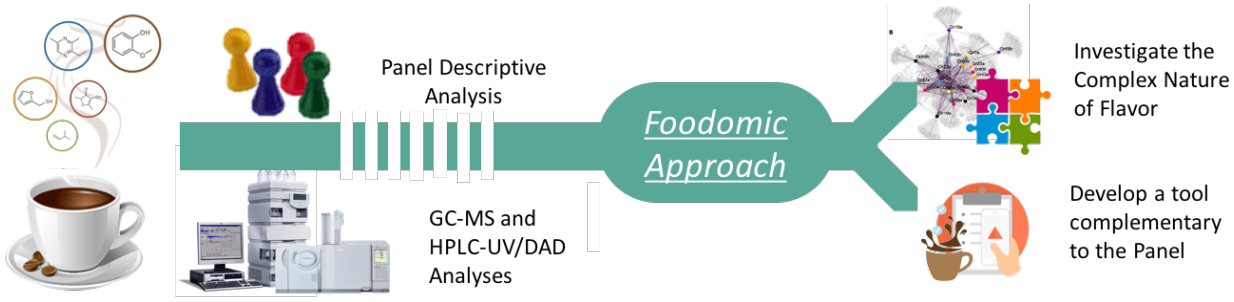
However, the lack of a direct relationship between the pool of chemicals used in the models and their sensory description stressed the need to support the sensometrically defined chemical fingerprints with an inter-approach validation. Molecular sensory science approach was therefore applied to investigate the pool of compounds able to characterize the *woody* and *flowery* notes. This compounds selection has been compared to that used to build the predictive models with a good consistency.

Within the pool of notes, some of them are known to be "taste attributes". In this step of the project, the non-volatile fraction, whose fingerprint was obtained by HPLC-UV/DAD analysis of the filtered brew, has been included in the modelling workflow to better understand its contribution in the flavour definition. The inclusion of the non-volatile fraction not only did not provide any significant improvement in the performance of the prediction models, but also made the whole workflow unsuitable for routine implementation.

In the last part of the project the developed models have been optimized and tested with a new samples selection trying to simulate their use in routine quality control. Last, all chemical fingerprints have been merged to develop a unique model able to make statistical inference on all

the sensory notes together with rather satisfactory results, although some compromises had to be accepted.

Core of the project has been the data elaboration by chemometric tools such as PCA, MFA, PLS-DA, parametric and non-parametric regressions.
The *–omic* approach together with chemometric tools are an important step forward compared to the conventional food flavour analysis because afford a better and more comprehensive investigation of the complex interactions behind the food flavour. If reliably and robustly developed, the chemical analysis can become a valid support for quality control purposes[3], although sensory panel contribution still continue to play an irreplaceable role.



References:

1. Gallo, M. & Ferranti, P. The evolution of analytical chemistry methods in foodomics. *J. Chromatogr. A* **1428,** 3–15 (2016).
2. García-Cañas, V., Simó, C., Herrero, M., Ibáñez, E. & Cifuentes, A. Present and Future Challenges in Food Analysis: Foodomics. *Anal. Chem.* **84,** 10150–10159 (2012).
3. Chambers IV, E. & Koppel, K. Associations of volatile compounds with sensory aroma and flavor: The complex nature of flavor. *Molecules* **18,** 4887–4905 (2013).

1. Introduction

1.1 Food Flavour: a complex, multisensory perception

Food, is going to move from being considered just a source of energy to a source of pleasure and well-being. Some but not all of human eating behaviours are indeed guided by basic homeostatic regulation. Other influences in this behaviour include cognition, emotion and reward and may lead to eating beyond, or below homeostasis, e.g. obesity or anorexia respectively.
Pleasure and reward mechanisms play a central role in the control of human food intake together with other factors such as genetics, circadian rhythms, reproductive status and social factors. Evidence for the complexity of eating behaviour can be found in the influence of all five primary sensory systems as well as the visceral sensory system and gut–brain interactions.[1]
Recent neurophysiological studies in particular made by Kringelbach and its group evidenced that the pleasure evoked by food is remarkably similar to that of other rewards, suggesting a unitary pleasure system, whether engaging with food, sex, social or higher-order rewards.

To understand physiological concept of pleasure, it is important to consider the main challenge for the brain which is to balance resource allocation successfully for survival and procreation. Different rewards compete for resources over time to achieve this balance. It can be useful to consider the typical cyclical time course shared between all rewards with distinct appetitive, consummation and satiety phases defined as Pleasure Cycle to understand the multi-faceted nature of pleasure, (Figure1.1.1A).
Many other neurological and tonic changes occur in the same time cycle described by the pleasure cycle and therefore the Figure 1.1.1A can be expanded into an elaborate multilevel model of food intake. (Figure 1.1.1). In addition, foods, with their peculiar composition, contribute actively and influence this circle.
The model links the pleasure cycle with the cyclical changes in hunger levels related to the initiation and termination of meals and the way food intake is ruled by the interaction of given signals from the body, e.g. from the oral cavity, brain, stomach and intestines etc…

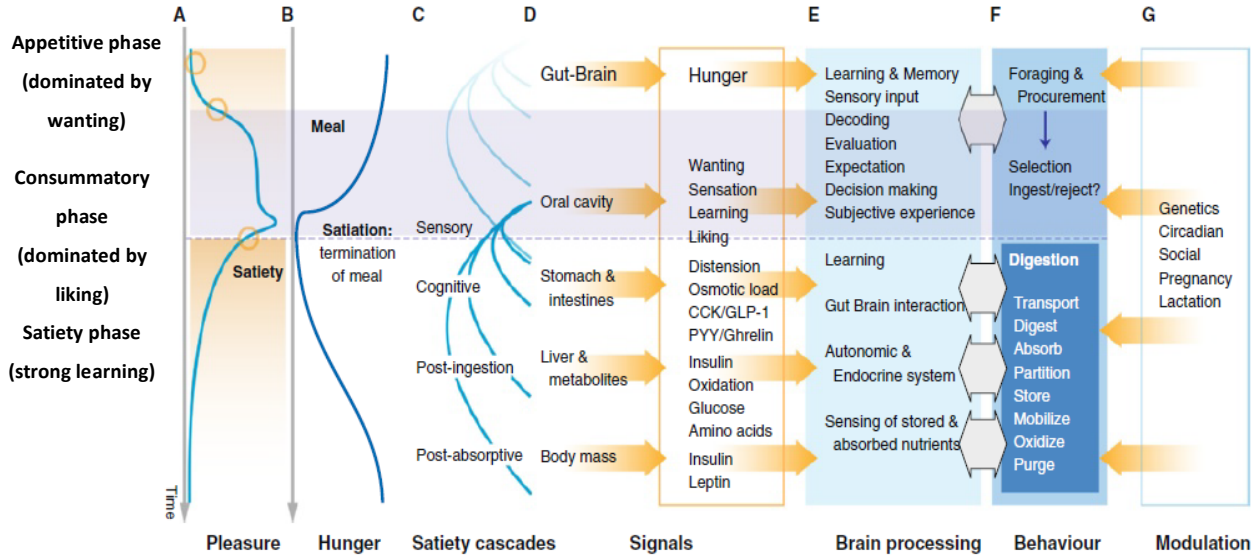


Figure 1.1.1 Multilevel model of food intake over time. The changes at each level before, during and after meals are shown in each column which summarises the episodic and tonic changes over time (moving from top to bottom): A) pleasure cycle, B) the levels of hunger, C) satiation/satiety cascade (sensory, cognitive, post- ingestion and post-absorptive signals), D) origin of signals (gut-brain, oral cavity, stomach and intestines, liver and metabolites and body mass) and signal carriers, E) brain processing, F) behavioural changes including digestive system and G) general modulatory factors 2

Food choice and intake are therefore influenced by each of these levels; but, a high portion of what we perceive concerning the food we assume, pass through the five classic senses (vision, hearing, smell, taste, and touch).[1–4]

Among this group, smell and taste are perhaps the two most important senses involved in eating nevertheless, food pleasure is not simply given by the sum of these stimuli but is something more: a cognitive perception defined *flavour*.

Taste and olfaction are mediated by molecules with different properties and by different physiological structures and are considered “chemical senses” because they convert chemical signals into action potentials in sensorial neuronal fibres (Figure 1.1.2)

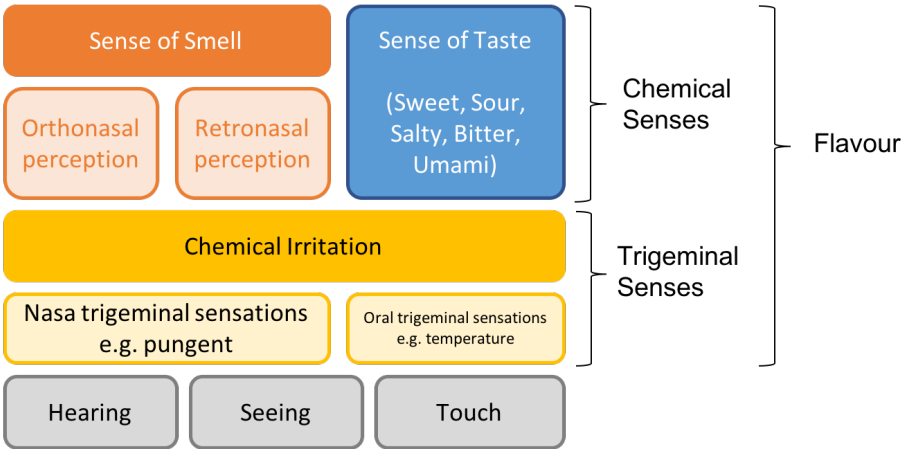


Figure 1.1.2 Scheme of the Flavour composition

Many scientific efforts have been made in the last years to understand better how all the different perceptions interact in the "construction" of flavours.
One possible strategy to decrypt the complex flavour perception is to classify its components in intrinsic and extrinsic factors[5]; intrinsic factors include taste, smell and part of the texture while aspects like food colours, sounds, packaging are extrinsic.
Food flavour is hence defined as a multisensory phenomenon, given by the integration of taste, olfactory, and other sensory (e.g. somatosensory) information into a perceived property of the food.[5]
It is therefore clear that the interactions between flavour components determine what we perceive and what we like in foods and drive the consumers' choices.

Sense of taste

The sense of taste is mediated by taste cells mainly located on the surface of the tongue and palate. The taste receptors are present on the apical region of these cells; they belong up to four morphological types: I (dark), II (light), III (intermediate) and IV (round).
Cells with different receptors perceive different tastes (e.g. type II allows the perception of sweet, bitter and umami tastes whereas type III cells perceive sour and salty tastes).

Groups from 50 to 100 cells aggregate to form taste buds; these buds present a pore where chemical compounds responsible for taste feelings can enter and interact with the receptors.
The transduction of the taste signals is different depending on the stimuli; direct ion channels mediates salt and sour, ion channels mediate for sour and bitter, while protein G receptors mediate for bitter, sweet and umami perceptions.
Taste signals are transmitted to the solitary tract in the brainstem and then retransmitted to the thalamus from where they reach the primary gustation cortex in the insula.

Taste active compounds present high molecular weight and a low/null volatility. Each of the five basic taste modalities is induced by chemical compounds with specific structures: (i) inorganic ions such as sodium, calcium, potassium are responsible for the salty perception, (ii) compounds able to induce a variation of the H^+ ions on the surface of the tongue are responsible for the sourness, while (iii) a wide range of chemical structures including poly-alcohols, aldehydes, ketones and aromatic molecules mediate for sweet, bitter and umami feelings.
A key factor about taste stimuli is that they elicit the most basic human emotions of pleasure (sweet) and disgust (bitter), which are not learned; they are hard-wired in the brainstem from birth.[6]

Sense of smell

Differently from the taste the olfaction is mediated by volatiles characterised by a low polarity and molecular weight (usually below 300 Dalton).
The odour perception starts when volatile compounds come into contact with the Odour Receptors (OR) expressed on olfactory sensory neuron (OSNs) cilia at the level of the olfactory epithelium (regio olfactoria).
The sense of smell can be splitted in two main perceptions:

- Orthonasal perception, when the odorants reach their receptors through the nostrils;
- Retronasal perception, when the odorants achieve the region olfactoria from the mouth cavity through the pharynx, i.e. food aroma is perceived during eating, and in particular during swallowing.

The odour perception is mediated by about 400 receptors (mainly belonging to the G protein family) able to recognise thousands of volatile compounds[7]. After binding, the sensorial olfactory nerve starts an action potential conducting information to the olfactory bulb (OB). Recently on the basis of psychophysical studied, Bushdid et al[8] indicated in around 10^{12} the number of human's detectable odorants; this huge number, combined to the limited number of OR shapes, suggests that at the olfaction implies that odours are detected by a combinatorial code mediated by the brain. Unlike taste, the neurological and cognitive responses to odour stimuli seemed mostly due to a learning process that can be the basis of the wide variability of odour perceptions and of the complexity behind this sensory stimulus[6].
At a physiological level, several characteristics distinguish olfaction from other sensory perceptions and makes it more complex and fascinating to be studied. It includes: i) the multiplicity of ORs, ii) the processing of the olfactory signal in the OB, iii) the direct projection to the cortex, and iv) the difficulty for humans to provide a verbal description of odour perceptions[9].

Aroma and taste deeply interact one another in the definition of flavours with a Cross-modal sensory integration.
This integration has been inferred from the influence of one modality on responses to another. Commonly, this is an enhanced (sometimes supra-additive) response to information from one sensory system due to concurrent input from another modality. Odour qualities, indeed, when added to tastants in a solution, can modify the taste intensity and vice-versa; this modification is often an enhancement as demonstrated by the use of strawberry or vanilla aroma compounds to enhance the sweetness of sucrose solutions [10,11].
A further but important consideration in the perspective of foodomic studies is the evidence that the odour-taste interaction can modify the perception of aroma and taste components present singularly in sub-threshold levels. Psychophysical studies detect that the sensitivity to benzaldehyde was significantly increased by the presence of a subthreshold concentration of saccharin in the mouth [12].
Integration of information from physiologically distinct sensory modalities is a general property of the mammalian nervous system.
Integration across sensory modalities is reflected in the presence of multimodal neurons that receive converging sensory information; these neurons may indeed respond specifically to combinations of different sensory inputs, or sensory-specific neurons may respond to modulations of other sensory pathways.
This neural enhancement is then behaviourally reflected in responses to multimodal stimuli. Like many other cross-modal sensory integrations also those aroma-taste depends on the temporal synchrony and moreover, the integration is stronger when both the stimuli are consistent one

another[13]. In order to form a new flavour, olfactory and taste information have to reach the receptorial structures together and at the same time.
Compared to other multimodal sensory pairs (e.g. audio-visual) those involving aroma and taste are less sensitive to onset discrepancies; this higher flexibility is a marker of a high adaptability of chemosensory binding.
In order to face this complexity several neurophysiological studies have been carried out at the beginning by considering taste and aroma separately and then putting them together.
In humans, the taste pathway ascends from the nucleus of the solitary tract in the brainstem to the hypothalamus and to the taste area of the somatosensory thalamus, and, from there, it extends to the primary taste cortex.
At the neocortical level, taste stimuli alone activate the primary taste centers in the tongue area and insula, whereas, as previously mentioned, odor stimuli by themselves activate the primary olfactory receiving area in the orbitofrontal region of prefrontal cortex.
When both taste and odor are stimulated together, an enhanced activity in several regions activated by the independent stimulations as well as additional activity in contiguous areas around the primary receiving areas is observed. This enhancement indicates that the flavour perception is more complex than that would occur through simple addition of taste and smell pathways.
Physiological studies suggest that odor (in particular) and taste are heavily conditioned by multi-sensory inputs and are heavily dependent on context and learning/rewards mechanisms.
Figure 1.1.3 reports the principal neurological pathways that lead the signals from the sensory receptors to the brain elaboration areas

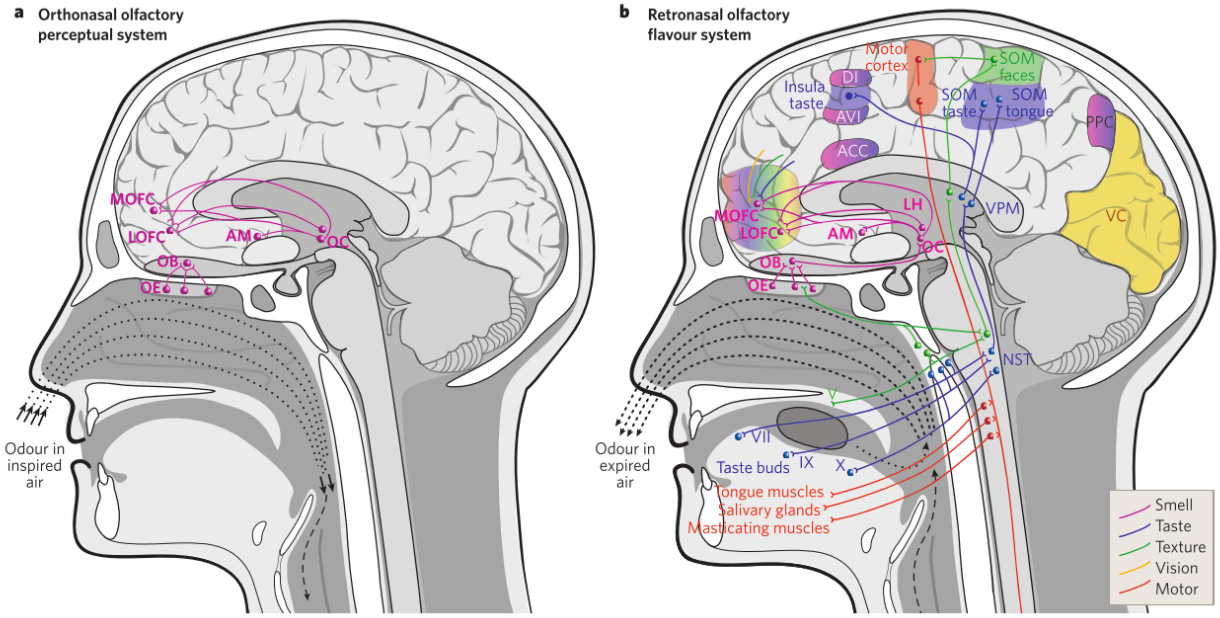


Figure 1.1.3 Brain systems involved on ortho and retronasal olfaction **a** and **b** ; blue and pink lines show the trasmission ways from the taste and odour receptors and the brain cortex. Green, Yellow and Red lines correspond to Texture, Vision and Motor pathways respectively. Legenda: ACC, accumbens; AM, amygdala; AVI, anterior ventral insular cortex; DI, dorsal insular cortex; LH, lateral hypothalamus; LOFC, lateral orbitofrontal cortex; MOFC, medial orbitofrontal cortex; NST, nucleus of the solitary tract; OB, olfactory bulb; OC, olfactory cortex; OE, olfactory epithelium; PPC, posterior parietal cortex; SOM, somatosensory cortex; V, VII, IX, X, cranial nerves; VC, primary visual cortex; VPM, ventral posteromedial thalamic nucleus. [6]

.

As reported in figure 1.1.4 the flavour perception involves many different food attributes to be collected and elaborated by a "sensory packaging centre" in the brain; that's why flavour is considered a cognitive perception[6].

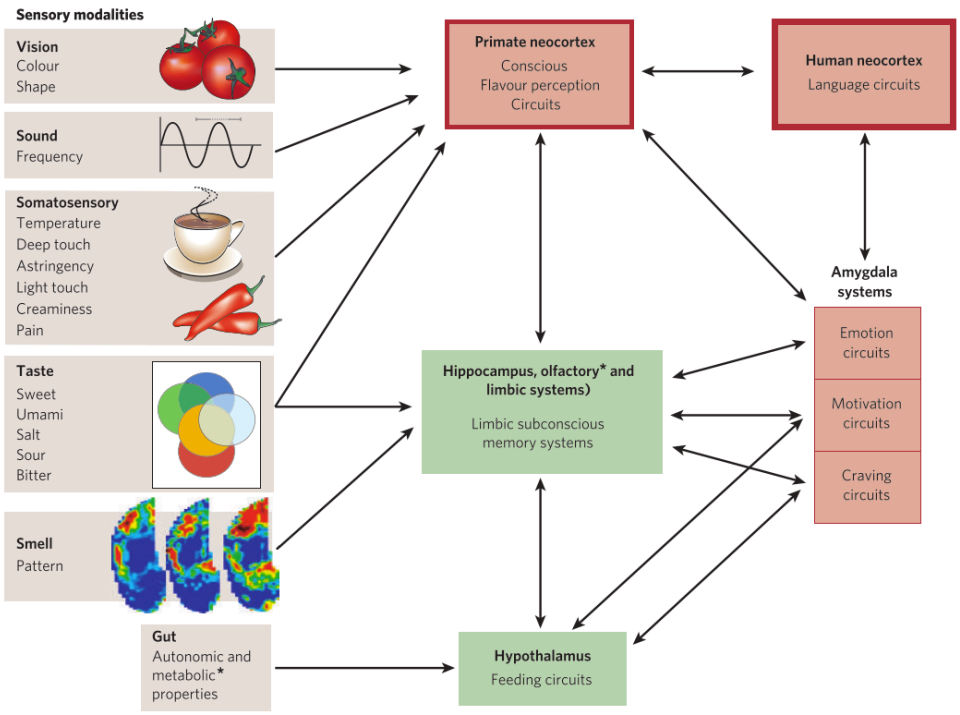


Figure 1.1.4 The human brain flavour system. The diagram shows the areas involved in the perceptual, emotional, memory-related, motivational and linguistic aspects of food evaluation mediated by flavour inputs. On the left all different food properties playing a role in food flavour definition red and green square reports the brain systems that mediate flavour perceptions. Red squares involved conscious perceptions while green ones are related to unconscious regulation[6].

The building of complex food flavours involves different modes of interaction within sensory modalities; the blending of odors to form completely new odors is defined as a *synthetic* interaction (similar to the blending of the light wavelengths) in which the final perception may not be linked to the odor descriptions of each single odorant. By contrast, the mixing of tastes is seen as an *analytic* process, because individual taste qualities do not fuse to form new qualities and can be distinguished from one another in the mixtures, similarly to what happens with single instruments of an orchestra during a concert.
The cognitive nature of flavour implies the introduction of a further new category of interactions, named *fusion,* in which different sensations (arising from olfaction and gustation senses) are combined to form a new perception rather than a new sensation like in *synthetic* interactions[14,15].

Food flavour can thus be intended as a functionally distinct sense which is cognitively "constructed" from the integration of distinct, physiologically defined sensory systems (mainly olfaction and taste) in order to perceive and identify objects that are important for our survival or pleasure [16,14].

FLAVOUR EVALUATION: A KEY STEP TO IMPROVE FOOD QUALITY

Food flavour, despite its complexity is one of the major determinants of food choice and usually represent a key feature in food selection [17].
Sensory aspects of foods are easily evaluable from the consumer and are essential for its satisfaction. Since flavour drives food quality, consumer acceptance and demand of evaluation and clear understanding of sensory aspects of foods are becoming of high importance for food industry[18]. The reported scheme[19] shows the main processes behind food choice and intake in which the perception of sensory properties is one of the pillars. The importance of the three main blocks can variate according to the food nature; in the case of comfort food, for example, the nutritional block is less important compared to the perception block[20] (Figure 1.1.5).

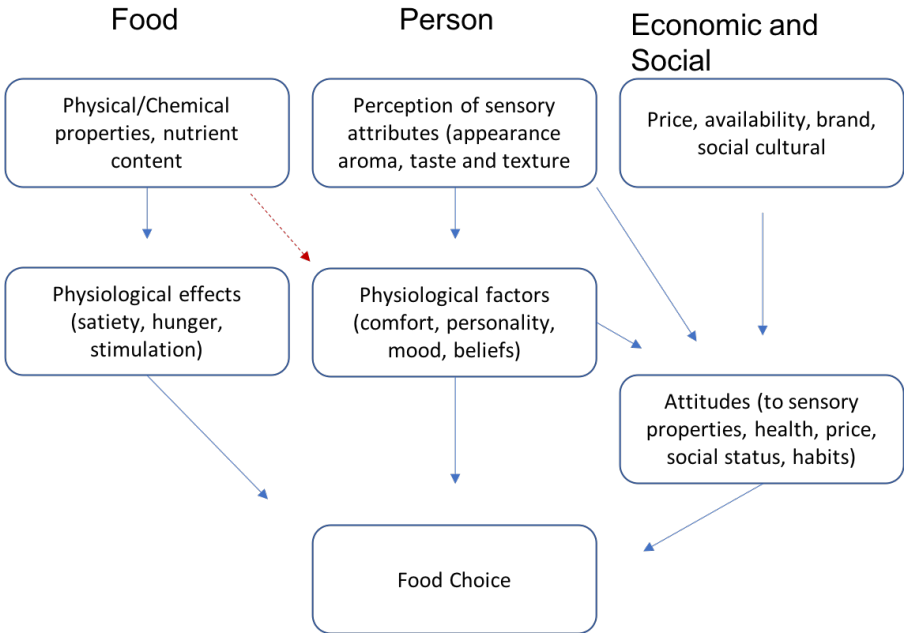


Figure 1.1.5 Scheme of the factors involved in food choice.

Many efforts have been performed by food industries to evaluate and control the sensory impact of their products. Thus food sensory evaluation has become a powerful tool for (1) lowering product costs, (2) recipe optimization, (3) innovation & product development, (4) quality monitoring and benchmarking, and (5) analysis of the "drivers or likeability"[18].

Food sensory evaluation

The field of sensory evaluation grew in parallel with the processed food industries. The food industries needed and still need to know the sensory properties of their products; this information is precious in many steps of the production chain (to evaluate the impact of some technological changes on the final product quality), in the development of new products according to consumers expectations and demands.[21]
Sensory evaluation is defined by Stone and Sidel[22] as "the scientific discipline used to evoke, measure, analyze, and interpret human reactions to those characteristics of foods and beverages as they are perceived by the senses of sight, smell, taste, touch, and hearing."

With this definition, the field of sensory analyses is intended as a pool of scientific methodologies that requires precise rules and protocols[23].
Sensory evaluation guidelines indeed provide instructions for samples selection, preparation and analysis. The second word included in the definition is “measure”; sensory evaluations is a quantitative science in which numerical data are collected to establish a well-defined relationship between products characteristics and human perception[24].
Likewise, in many other quantitative science, one essential part of the approach concerns the data elaboration and this statement is even more true for sensory evaluation.
The human responses, and the sensory tests, include many sources of variation and most of them cannot be controlled and limited by a standardized and meticulous analytical protocol.
Examples include the mood and motivation of the participants, their innate physiological sensitivity to sensory stimulation, and their past history and familiarity with similar products.
At the end of the food sensory analysis the data interpretation is important to change data into results; it is therefore important to look everything within the context, the background and the objectives that has to be achieved.
Nowadays, most sensory testing is performed in an industrial environment as a support for business concerns and strategic decisions. Sensory analyses are often used as a tool to reduce the risk and uncertainty in decision making.
In addition, sensory testing is of particular importance in the development of new food product where the final perception of the product has to be checked.
Beyond their importance in product development sensory evaluation may provide information to many different parts of the “food life”, from the packaging to the shelf-life.
In the recent years the quality control program of many foods and comfort foods in particular, include a sensory evaluation step. The current sensory evaluation methods comprise a set of measurement techniques with established protocols and rules used both in industry and in academic research.

One of the primary issues about sensory testing concerns the selection of the suitable test; sensory test can be classified in three groups according to their primary purpose and most valid use. Each group has a different goal and is carried out with participants selected with different criteria:

- *Difference testing*: these are the simplest sensory test ad they aim to answer whether any perceptible difference exist between two types of product. A classic example of this category is provided by triangle test, where two products were from the same batch while the third one was different. Judges would be asked to pick the odd sample among the three.
- *Descriptive analyses*: in this category, the perceived intensities of the sensory characteristics of a product are quantitatively determined. One of the most important methods is termed Quantitative Descriptive Analysis® or QDA® that merges the completeness of the human perception with the modern experimental design techniques.
  Descriptive analysis has proven to be the most comprehensive and informative sensory evaluation tool. It is applicable to the characterization of a wide variety of product changes and research questions in food product development. The information can be related to consumer acceptance information and to instrumental measures by means of statistical techniques such as regression and correlation.

- *Novel methods*: these methods have been developed as alternatives to classical descriptive analysis with the advantage of flexibility, fastness and possibility to be performed by experts, trained assessors or even by untrained consumers.
  First example of these new sensory methods is the "Sorting Task"; in this test, the assessors have to sort the samples according to their similarities within the parameters under evaluation.
  The main inconvenience of this test is that it is based on similarities between samples and the sensory attributes remains in the background. Flash profiling is a combination of free choice profiling (FCP) and ranking methods that try to merge the power of a descriptive method and the speed of these novel techniques. Flash profiling consist of two steps; in the first the assessors are asked to observe, smell or taste the samples and to generate a set of attributes able to discriminate the samples while in the second steps the assessors have to rank the samples from the least to the most on each of their attributes.
  Check All That Apply (CATA) technique was applied to understand consumer's preferences; it provides easy to read and interpret data maps. Assessors have to select among those provided with a questionnaire the attributes (or the phrases) suitable for defining the sample in test.
  The evaluation of the data obtained from these new sensory tests requires the application or multivariate statistical analyses such as the general procrustes analysis of the multiple factorial analysis.
  In addition, since these techniques are based on comparative methods (they make the sensory analysis possible also for untrained assessors), the use of a stable reference within the test is often needed.

Today most of food consumption is linked to the pleasure perceived during its intake and therefore the evaluation of food sensory impact became an integral part of the quality control (QC) and quality assurance. Sensory tests can proficiently be used in the sensory assessment of the incoming raw material, semi-finished products final product quality monitoring but can also be employed in routine programs for shelf-life assessment.

Due to the time and the resources needed to assemble and train the panel, prepare the samples for testing, analyse and report sensory data, it can be quite challenging to apply sensory techniques to quality control as an-online assessment. For this reason, one of the needs of modern comfort (but not only) food industries is to establish a relationship between sensory response and instrumental measures. If it is correctly performed, the instrumental measure can be complementary to sensory data and in some (rare) cases, can even substitute the human contribution. One of these possible scenarios is represented by those conditions that requires a rapid turnaround not always in compliance with human rhythms. Substitution of instrumental measurements for sensory data may also be useful if the evaluations are likely to be fatiguing to the senses, repetitive, involve risk in repeated evaluations (e.g. insecticide fragrances).

References:

1. Kringelbach, M. L., Stein, A. & van Hartevelt, T. J. The functional human neuroanatomy of food pleasure cycles. *Physiol. Behav.* **106,** 307–316 (2012).
2. Kringelbach, M. L. The pleasure of food: underlying brain mechanisms of eating and other pleasures. *Flavour* **4,** 20 (2015).
3. Berridge, K. C. & Kringelbach, M. L. Pleasure Systems in the Brain. *Neuron* **86,** 646–664 (2015).
4. Kringelbach, M. . Food for thought: hedonic experience beyond homeostasis in the human brain. *Neuroscience* **126,** 807–819 (2004).
5. Prescott, J. Multisensory processes in flavour perception and their influence on food choice. *Curr. Opin. Food Sci.* **3,** 47–52 (2015).
6. Shepherd, G. M. Smell images and the flavour system in the human brain. *Nature* **444,** 316–321 (2006).
7. Audouze, K. *et al.* Identification of Odorant-Receptor Interactions by Global Mapping of the Human Odorome. *PLoS One* **9,** e93037 (2014).
8. Bushdid, C., Magnasco, M. O., Vosshall, L. B., Keller, A. & Mixture, M. Humans Can Discriminate More than 1 Trillion Olfactory Stimuli. *Science (80-.).* **343,** 1370–1373 (2016).
9. Tromelin, A. Odour perception: A review of an intricate signalling pathway. *Flavour Fragr. J.* **31,** 107–119 (2016).
10. Frank, R. A., Ducheny, K. & Mize, S. J. S. Strawberry odor, but not red color, enhances the sweetness of sucrose solutions. *Chem. Senses* **14,** 371–377 (1989).
11. Frank, R. A. & Byram, J. Taste–smell interactions are tastant and odorant dependent. *Chem. Senses* **13,** 445–455 (1988).
12. Dalton, P., Doolittle, N., Nagata, H. & Breslin, P. a. The merging of the senses: Integration of subthreshold taste and smell. *Nat. Neurosci.* **3,** 431–432 (2000).
13. Spence, C. Multisensory flavour perception. *Curr. Biol.* **23,** 373–394 (2013).
14. Prescott, J. Flavour as a psychological construct: implications for perceiving and measuring the sensory qualities of foods. *Food Qual. Prefer.* **10,** 349–356 (1999).
15. Prescott, J. Multimodal Chemosensory Interactions and Perception of Flavor. in *The Neural Bases of Multisensory Processes* (eds. Murray, M. & Wallace, M.) (Taylor & Francis, 2012).
16. Gibson, J. J. *The senses considered as perceptual systems*. (Praeger, 1966).
17. Clark, J. E. Taste and flavour: their importance in food choice and acceptance. *Proc. Nutr. Soc.* **57,** 639–643 (1998).
18. Regueiro, J., Negreira, N. & Simal-Gándara, J. Challenges in Relating Concentrations of Aromas and Tastes with Flavour Features of Foods. *Crit. Rev. Food Sci. Nutr.* **8398,** 0 (2015).
19. Shepherd, D. Dietary Salt Intake. *Nutr. Food Sci.* **85,** 10–11 (1985).
20. Spence, C. Comfort food: A review. *Int. J. Gastron. Food Sci.* **9,** 105–109 (2017).
21. Lawless, H. T. & Heymann, H. *Sensory Evaluation of Food*. *Science* (Springer New York, 2010).
22. Stone, H. & Sidel, J. L. *Sensory evaluation practices*. (Elsevier Academic Press, 2004).
23. Murray, J. ., Delahunty, C. . & Baxter, I. . Descriptive sensory analysis: past, present and future. *Food Res. Int.* **34,** 461–471 (2001).
24. Britta Folmer. The Craft and Science of Coffee. in *The Craft and Science of Coffee* (eds. Blank, I., Farah, A., Giuliano, P., Sanders, D. & Wille, C.) iii (Elsevier, 2017).

1.2 Comfort foods and coffee in modern dietary scenario

Food and eating play an increasingly central role in individual lives. Food science, over the last fifty years, has been focused on issues concerning food supply, safety and nutrition but at the present, and in our society, these needs are adequately satisfied by a lot of manufacturers; then other issues are coming out[1].

Today, one of the main aspects of food consumption is that consumers wants to enjoy their food.

Many foods of our daily diet are chosen and consumed because of the pleasure perceived during their intake.

This wide and heterogeneous class of foods are known as "*comfort foods*". Over the years these foods have been defined in many ways; Wansink and Sangerman offered a general and versatile definition: "a comfort food is a specific food consumed under a specific situation to obtain psychological comfort". A similar definition has also been provided by Locher defining a comfort food as "any food consumed by individuals, often during periods of stress, that evokes positive emotions and is associated to significant social relationships".

Studies on comfort foods classifies this category of foods in three main groups:

Nostalgic Foods: Comfort foods whose consumptions evokes in the consumer a sense of familiarity. These foods are chosen according to the experience of the people who make the choice, they are linked to particular memories like events of their childhood that make them feel better.

Convenience Foods: When people seek comfort, the need is immediate; foods able to provide effortless gratification can thus be considered comfort foods. Generally, this group contains foods like chocolate, ready to eat frozen foods, home-delivered pizza, potato chips etc… The convenience foods the comfort does not come only from the food itself but also by its immediate availability.

Physical Comfort Foods: this group includes foods that offer comfort either through their physical attributes or the changes they make on the consumer's body. Examples of these physical comfort foods are alcoholic drinks and coffee.

Although food consumption has always been linked to the pleasure perceived during food intake, only in the last decade the notion of comfort foods begun appearing regularly in popular magazines, televisions, literature and advertising. This increase of interest might be due to the fact that the daily life in modern world is even more stressful, competitive and put the people more under pressure than in the past so that they need the positive emotions given by comfort foods[1–3,4].

COFFEE, ONE OF THE MOST IMPORTANT COMFORT FOODS CONSUMED WORLDWIDE

Many people all over the world consider coffee as a comfort food and drink it because of the pleasure linked to its peculiar flavour. Coffee flavour is the final expression of a long chain of cultivation and transformation that links the seed to the cup. This long journey has been summarized by Yeretzian in the Handbook of Odour (ed. Springer)[5] in three key factors: *Predisposition*, *Transformation* and *Consumption*.
Within this definition, the consumption is considered a real part of coffee value creation chain because the experience offered by coffee consumption is strictly related to how the final consumers prepares and drinks the beverage.
The Predisposition considers all the pre-harvesting properties of the coffee; each variety is genetically characterized by a specific pool of aroma precursors that origin different aroma signatures.
The transformation regards all the steps from the harvested bean up to the final roasted product; these processes are essential to obtain a high-quality product.
Thanks to its worldwide diffusion, coffee is the most important (in financial terms) agricultural commodity after petroleum.
According to the 2015-2016 annual review of the International Coffee Organization (ICO) the coffee consumption has continued to increase reaching a record of 151.3 millions of bags consumed worldwide; the main player of this increase is the Asia & Oceania market that registered a 3.75% of increment (ICO 2016). Europe is the principal coffee consuming continent with a consumption around 50 million bags.

As a consequence of this ever-increasing demand of coffee, the need of quality assessment and control protocols is increasing as well.
The acquisition of a deep knowledge about their products is becoming mandatory for coffee industry, not only on the final product, and/or at the end of the production chain, but also at its beginning. The quality assessment and control of the raw material is crucial to design and produce coffees in line with consumers desires and expectations.

Coffee quality assessment involves many different parameters that are summarized in figure 1.2.1
Among all these parameters, aroma and taste are considered as coffee primary quality attributes and their evaluation is mandatory for the product quality definition at each step of the production chain.

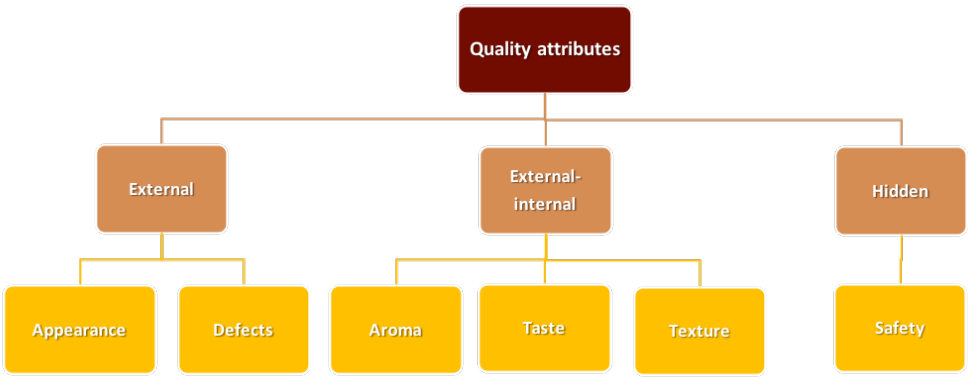


Figure 1.2.1 Food properties important in its quality definition (UN. (2007). Safety and quality of fresh fruit and vegetables: A training manual for trainers. United Nations. Retrieved from http://www.unctad.org/en/docs/ditccom200616_en.pdf (Accessed 15/02/2015).)

Despite the huge number of researches in this field, the chemistry behind the relationship between coffee flavour and the elicited sensory perception is still a challenge from a scientific and industrial point of view due to the complexity of this fascinating attribute[5–10].
Nowadays, human trained panels make the evaluation of the sensory properties of a good coffee. Sensory analyses have been the object of a lot of work oriented to the standardization and the objectivation of coffee evaluations, although they still are affected by panel's subjectivity and by being time-consuming in terms of panel training, alignment and number of assessments per day.
This PhD thesis deals on a multidisciplinary study on the relationship between the chemicals of the coffee's flavour and its sensory properties.

Coffee come from the botanical genus *Coffea* belonging to the family of *Rubiaceae.*
Among the 70 species of *Coffea* genus, only three are cultivated: *Coffea arabica*, *Coffea canephora* and *Coffea liberica* which cover respectively 75%, 25% and less than 1% of the market share.
The coffee shrub grows in high tropical altitudes (600–1200m) with an annual average temperature of 15–25 °C. They need moderate moisture and cloudiness.
The plant produces up to 40 years but the maximum yield is obtained after 10–15 years[11].

The fruits of the coffee shrub, or berries have a green skin that becomes a deep red when the ripening is complete. A scheme of the morphology if a coffee berry is reported in figure 1.2.2.

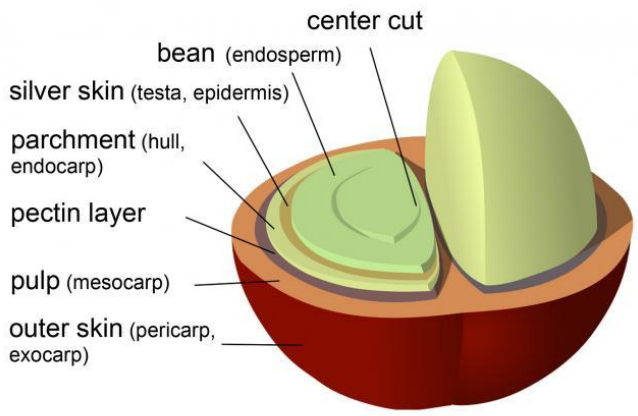


Figure 1.2.2 Scheme of a coffee bean Source: www.feedipedia.org

After the outer skin (pericarp), the berry consists of the sweet mesocarp (or pulp) and more in depth the real coffee beans (the endocarp or parchment). The beans are two elliptical hemispheres with flatten adjacent sides. The beans are singularly covered by a thin spermoderm layer called silver skin. Each berry usually contains two beans, but in the 10-15% of the cases some berries contains just one single bean; in these cases, it is called "peaberry" or "caracol" and it is sold at a higher price.
Coffee harvesting occurs from December to February and from May to August according to the geographical region and can be done by hand-picking (more expensive and accurate) or by strip-picking (cheaper but less accurate).
The main difference between these two techniques is that by hand-picking just the full-ripe berries are harvested while in the strip-picking all the berries from the brunch are collected.
Processing started with the removing of the pulp that can be by a dry of wet procedure: the dry protocol involved the spreading of the whole fruit on sun-drying terraces. Coffee beans are

collected after the shrinking of the pulp. The wet procedure generally leads to a better-quality coffee and found application for Arabica coffees.
Fresh berries are squeezed by a rotating cylinder; during the squeezing the seeds are detached from the pulp without any damage. The mucilaginous layer that surrounds the beans is then removed by enzymes during the 12-48h of fermentation. Fermented beans covered by the silverskin are dried (naturally or by machines) and dehulled before being shipped.
Green beans are heat treated in a process called roasting that is the main responsible of the formation of the typical coffee aroma.
The roasting procedure is characterized by four steps that changes the chemical composition of the coffee beans: (i) the drying; at almost 50°C the proteins denature and the water evaporates, then (ii) the development occurs at 100/150°C when the browning starts and volatiles are released (water, CO_2 and CO), (iii) the decomposition; at 180°C/200°C the beans pop and burst with a release of smoke and coffee aroma, finally (iv) the full roasting status is reached when the total amount of moisture drops at 1.5-3.5%.
Roasting is a very important step of the whole coffee production chain and it is fundamental to obtain a high-quality coffee aroma. For this reason, despite it is electronically controlled, it requires experience to be correctly carried out.

Non-volatile coffee components and their impact on the flavour

The final coffee flavour have its origin in the green raw beans. The chemical composition of raw green beans in the raw material on which the roasting process acts, is at the origin of the pool of volatiles of coffee aroma[6,12]. Figure 1.2.3 reports an overview of the chemical composition of green and roasted Arabica and Robusta coffee. Labels reports the % on the dry matter.

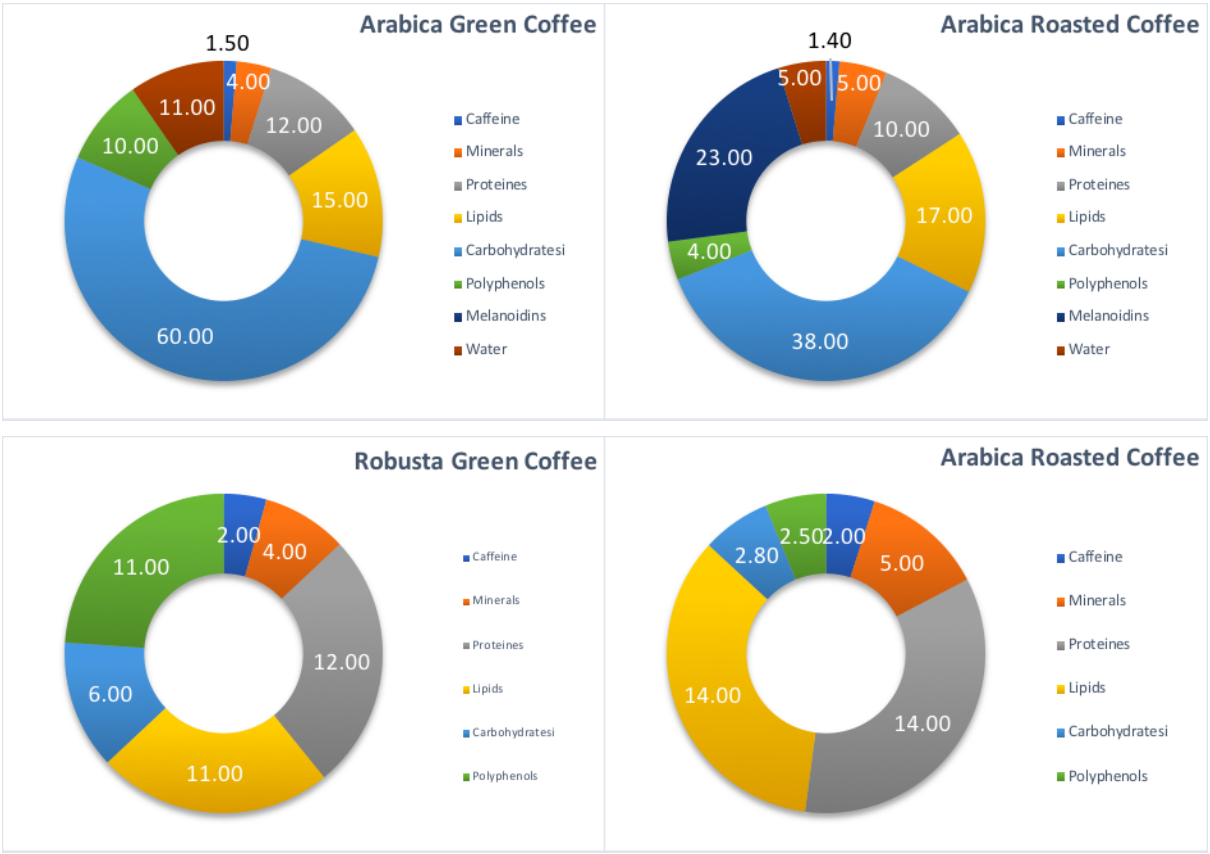


Figure 1.2.3 percentage composition of both Arabica and Robusta coffees. Both green and roasted beans have been considered.

Carbohydrates are products of the photosynthesis in plants and are the main portion of the coffee bean. By far the polysaccharides form the largest part of the green coffee and among them the main ones are mannan (gaclatomannan) and arabinogalactan fractions.
Polysaccharides are degraded and depolymerized during the roasting process where they play an important role as precursors in flavour generation.
In roasted coffee, polysaccharides play an important role in retaining volatiles and contribute to the perceived viscosity of the coffee brew. The amounts of sucrose in mature beans contribute to the coffee sweet note

Melanoidins are a heterogeneous group of nitrogen-containing colored polysaccharides; they can be at both high and low molecular weight but in coffee, the high molecular weight (HMW) fraction is prevalent (59%).
Melanoidins are of the particular interest both from a scientific and industrial view points since they are not chemically fully characterized but they confer the brown color to the brews that is intimately associated in consumers' minds to a high-grade product of desirable texture and flavour. Melanoidins are formed during the roasting process by cyclization, dehydrations, retro-aldolization, rearrangement, isomerization, and condensation of low molecular weight products of the Maillard reaction. They showed antioxidant properties in many *in vitro* and *in vivo* studies mainly based on a radical scavenging activity and on the metal chelating capacity[13].

Amino acids are present mainly as constituents of peptides and proteins, but free amino acids are also important on the coffee flavour formation.
The free amino acid fraction is largely transformed during the roasting by the Maillard reaction:

- sulfurated amino acids, cysteine and methionine, in green coffee are mostly bound in proteins, during roasting these amino acids react with reducing sugars in the Maillard reaction to form a pool of aroma active compounds as furfurylthiol, thiophenes and thiazoles.
- the reaction of hydroxy-amino acids (serine and threonine) with sucrose produce volatile heterocyclic compounds such as alkylpyrazines,
- pyrroles, pyrrolizidines, pyridines but also alkyl-, acyl- an furfurylpyrroles originate from the reaction of the Maillard intermediates with proline and hydroxyproline.

The proteic content of green coffee is around 10-13%. Roasting have an impact on the protein content that is fragmented, polymerized and integrated in melanoidins.
Among the nitrogenous compounds caffeine and trigonelline are very important.
These secondary metabolites are alkaloids whose biosynthesis takes place in the leafs or in the pericarp. Caffeine is produced by the plant for defense against herbivory, mollusks, insects, fungi and bacteria. The caffeine content in coffee is not affected by post-harvest process and roasting, it is extracted during the brewing procedure and assumed during coffee drinking. After consumption, caffeine is absorbed from the gastrointestinal tract and within 1h it is distributed in the body; it carries out its physiological stimulating effects after passing the blood-brain barrier.
Although caffeine as a pure chemical is assigned of a marked bitter taste (so that is used as a bitterness reference), its role in coffee bitterness is not so important.
Contrary to caffeine, trigonelline is rapidly degraded during roasting producing nicotinic acid, methyl-pyrazines, pyridines and pyrroles. Because of its reactivity, trigonelline has an impact on the overall aromatic perception of coffee and can thus be considered as an aroma precursor.

Lipids are present, in green coffee mainly at the level of the beans endosperm as coffee oils; therefore, some lipids are also present as waxes outside the bean. Lipids in coffee are important as carriers of flavours and contribute to the perceived texture and mouthfeel in the brew.
The lipid fraction consists of several classes of compounds: triacylglycerols, diterpenes and diterpenes esters, steroids, tocopherols and coffee wax. This fraction is responsible of the crema emulsion in espresso coffee preparation that carries flavour volatiles.

Acidity in an important attribute in coffee flavour and is often conversely associated to sweetness. The acid content of green coffee is around the 11% and is given by citric, malic, chlorogenic and quinic acids. This percentage decreases to 6% in roasted coffee because the degradation of these compounds to form volatile aroma compounds such as guaiacole and 4-vinylguaiacole. The heat treatment lead to the formation of formic, acetic, glycolic and lactic acids.

Chlorogenic Acids are present in the coffee beans in relatively high quantities. This group of chemicals comes from the conjugation of quinic acid (tetrahydroxy-cyclohexane carboxylic acid) with caffeic acid (3,4-dihydroxycinnamic acid). This conjugation originates a wide pool of isomers that makes the chlorogenic acids so numerous. The most common Chlorogenic acid in coffee is 5-*O*-caffeoyl-quinic acid (5-CQA).
Chlorogenic acids are biosynthesized in the perisperm and accumulated in the beans' endosperm; their content is progressively reduced during roasting. One of the transformation occurring during the heat treatment is the formation of the Chlorogenic lactones in the quinic portion of CQA. These lactones are characterized by bitterness and possible biological effects[14].
The biological effects of chlorogenic acids are under study, in particular they have been assigned of a potential antioxidant activity and a protective effect against degenerative diseases such as cardiovascular disease, cancers and diabetes II. In sensorial terms, these secondary metabolites contribute to bitterness and astringency.

Coffee volatiles and impact on coffee flavour

Aroma is one of the reasons (if not the main) for coffee consumption as a comfort food. Over the years, coffee aroma has been of interest for both its complexity and the role it has in the coffee value chain and in the dynamics behind coffee consumption. Volatile Organic Compounds (VOCs), responsible for the classical coffee aroma are arguably the most important quality-determinant of coffee. The pool of volatiles identified in the coffee aroma counts around 1000 compounds that range in concentration from ppm (part per million) to ppt (part per trillion). Among these compounds the studies on molecular sensory science pointed out a group of about 30 components defined as *coffee key aroma compounds*. A suitable mixing of these compounds in agreement with their odor activity values (OAV) affords to reproduce the well-known aroma of coffee[15] quite accurately.
Despite these important findings a question is still open: is the coffee aroma so simple? How the other 970 compounds impact on the different modulation of the overall coffee aroma? It is because of this question that the studies on coffee aroma are still a fascinating field of research[16].
As previously discussed, coffee volatiles derived from a large number of precursors present in the green coffee beans that are submitted to a series of reactions during coffee processing. The main chemical reactions occurring during roasting include Maillard reactions, phenolic acid and carotenoids degradation, Strecker degradation, breakdown of sulphur amino acids, hydroxyl-

amino acids, proline and hydroxyproline, degradation of trigonelline, chlorogenic acids and quinic acids and so on…

Furans are one of the most important class of volatiles from a quantitative point of view. They originate from the thermal degradation of carbohydrates, ascorbic acid and unsatured fatty acids during roasting. A summary of the most important furans is reported in table 1.2.1.

Table 1.2.1 Important Furans for coffee aroma

VOC	*CAS n°*	*Aroma Description*	*Sensory Threshold (ppb)*
Furfural	98-01-1	sweet woody almond fragrant baked bread	280
2-((methylthio)methyl)furan	1438-91-1	onion garlic sulfury pungent vegetable horseradish	-
2-furanemethanol acetate	623-17-6	sweet fruity banana horseradish	-
2-methylfuran	534-22-5	ethereal acetone chocolate	-
5-methyl-2-furancarboxyaldehyde	620-02-0	Sweet, caramellic, bready, brown, coffee-like	6000
furfurylformate	13493-97-5	ethereal	-
furfurylmethyl ether	6270-56-0	sweet spicy	-
furfuryldisulfide	57500-00-2	Roasted coffee, sulfurous cooked meat and liver, onion and garlic nuances. Slight eggy undertones	-

Furans are characterized by a general malty/sweet aroma perceptible in relative high odor threshold if compared to those of other odorants, but they are important for coffee aroma because they are present in high concentrations (3-115 ppm).

Pyrazines are the second most abundant class of chemicals present in high concentration combined and with a low odor threshold.
Pyrazines are mainly formed during roasting from Maillard reactions between amino acids and sugars[17], they are characterized by nutty, earthy, roasty aromas. Among this group the most important pyrazines are 2-methoxy-3-isopropylpyrazine, 2-ethyl-3,5-dimethylpyrazine and 2,3-diethyl-5-methylpyrazine.

Table 1.2.2 Important pyrazines in coffee aroma definition

VOC	*CAS n°*	*Aroma Description*	*Sensory Threshold (ppb)*
2,3-dimethylpyrazine	5910-89-4	Musty, nut skins, cocoa powdery and roasted with potato and coffee nuances	800
2,5-dimethylpyrazine	123-32-0	Nutty, peanut, musty, earthy, powdery and slightly roasted with a cocoa powder nuance	80
2,3-diethyl-5-methylpyrazine**	18138-04-0	Musty, nut skin, earthy, toasted, potato bin, green and meaty	0.09
2-ethenyl-3,5-dimethylpyrazine	157615-33-3	earthy musty	0.000012
2-ethenyl-3-ethyl-5-methylpyrazine	181589-32-2	earthy	0.000014
2-ethyl-3,5-dimethylpyrazine**	13925-07-0	Peanut, nutty, caramel, coffee, musty, cocoa, pyrazine and roasted	0.04
2-ethyl-3,6-dimethylpyrazine	27043-05-6	nutty	8.6
2-methoxy-3,5-	92508-08-2	earthy	0.006

VOC	*CAS n°*	*Aroma Description*	*Sensory Threshold (ppb)*
dimethylpyrazine			
2-methoxy-3,2-methylpropylpyrazine	24683-00-9	green pea green bell pepper green pea galbanum	-
2-methoxy-3-isopropylpyrazine	25773-40-4	pea earthy beany chocolate nutty	0.002
3-ethenyl-2-ethyl-5-methylpyrazine	181589-32-2	earthy	-
6,7-dihydro-5-methyl-5H-cyclopentapyrazine	23747-48-0	Sweet, nutty, roasted, toasted, grainy, coffee and corn with savory, meaty nuances	-
ethylpyrazine	13925-00-3	Nutty, musty, fermented, coffee, roasted,cocoa and meaty nuances	4000

Thiols and sulphur compounds are qualitatively considered among the most significant contributors to coffee flavour. This importance is due to their sensory potency.
Table 1.2.3 reports the main sulphur compounds described in the literature as coffee aroma components.

Table 1.2.3 Sulphur compounds important for coffee aroma

VOC	*CAS n°*	*Aroma Description*	*Sensory Threshold (ppb)*
Sulphur-containing compounds			
dimethyl trisulfide	3658-80-8	sulfureous, alliaceous, cooked, savory, meaty, eggy and vegetative with a fresh, green, onion topnote	0.001
bis(2-methyl-3-furyl)disulphide**	28588-75-2	Sulfurous, strong roasted meaty with a note reminiscent of sulfurol, savory, au jus and chicken nuances	0.00076
Methional **	3268-49-3	Vegetable oil. Creamy tomato, potato skin and French fry, yeasty, bready, limburger cheese with a savory meaty brothy nuance	0.2
Thiols			
3-mercapto-3-methylbutylformate	50746-10-6	sulfurous catty caramelised onion; roast coffee, roast meat, with a tropical nuance on dilution	0.0035
2-furfurylthiol **	98-02-2	Roasted coffee, sulfurous, with a burnt match note. It is savory meaty with chicken and fried onion nuances.	0.01
2-methyl-3-furanthiol **	28588-74-1	Sulfureous, meaty, fishy, metallic and roasted chicken-like	0.007
3-mercapto-3-methylbutylacetate	50746-09-3	roasty, fruity sulfurous sweet	-
3-methyl-2-butene-1-thiol **	5287-45-6	sulfurous smoke leek onion skunky	0.0003
methanethiol	74-93-1	Vegetable oil, alliaceous, eggy, creamy with savory nuances	0.02
Thiophene			
3-methylthiophene	616-44-4	fatty winey	-
Thiazole			
2,4-dimethyl-5-ethylthiazole **	38205-61-7	nutty roasted meaty earthy	-

Among the pool of sulphur compounds reported in table 1.2.3 those with ** are also part of the pool of coffee key aroma compounds defined by Grosch et al.[15]. These compounds are characterized by roasted (coffee-like), sulfury, meaty notes.

Furanones are mainly responsible for sweet caramel notes. This class of compounds originate from the Maillard reaction followed by aldol condensation.

Table 1.2.4 Furanones important for coffee aroma. Compounds marked with ** have been selected as coffee key aroma compounds thanks to their Odour Activity Values (OAVs) by Grosch et al 10.

VOC	*CAS n°*	*Aroma Description*	*Sensory Threshold (ppb)*
dihydro-2-methyl-3(2H)-furanone	3188-00-9	Sweet and solvent-like with a brown, rummy and nut-like nuance	0.005
2-ethyl-4-hydroxy-5-methyl-3(2H)-furanone	27538-09-6	Sweet, caramellic, bready, maple, brown sugar with burnt undernotes	20
3-hydroxy-4,5-dimethyl-2(5H)-furanone (sotolone)**	28664-35-9	extremely sweet strong caramel maple burnt sugar coffee	20
4-hydroxy-2,5-dimethyl-3(2H)-furanone (furaneol)**	3658-77-3	Sweet, slightly burnt brown caramellic, cotton candy with a savory nuance	10
5-ethyl-3-hydroxy-4-methyl-2(5H)-furanone (abhexon)**	698-10-2	Sweet, fruity, sweet caramel with a maple like nuance	7.5
5-ethyl-4-hydroxy-2-methyl-3(2H)-furanone	27538-09-6	Sweet, caramellic, bready, maple, brown sugar with burnt undernotes	1.15

Phenolic compounds are the last group of the main coffee aroma compounds.
These compounds originate from the thermal degradation of the chlorogenic acids (ferulic, caffeic and quinic acids) and their concentration after roasting is proportional to that of their precursors in the green beans.
The concentration of phenolic compounds is related to the species (Robusta is richer than Arabica) and to the geographical origin; it can range between 3 to 56 ppm. Phenolic compounds are characterized by a spicy sensory note; their smell is so peculiar to be defined as “phenolic”. Vanillin diverge from the guaiacoles sensory description and elicit a sweet sensation that can be defined as “vanilla-like”.
Table 1.2.5 reports the main phenolic compounds defining coffee aroma.

Table 1.2.5 Phenolic compounds important in coffee aroma

VOC	*CAS n°*	*Aroma Description*	*Sensory Threshold (ppb)*
Guaiacol**	93-51-6	Sweet, candy, spice, eugenol, vanilla, leather, spicy, smoky	2.5
4-ethylguaiacol **	2785-89-9	picy and clove-like with medicinal, woody and sweet vanilla nuances	25
4-vinylguaiacol **	7786-61-0	dry woody fresh amber cedar roasted peanut	0.75
Vanillin **	121-33-5	Sweet, vanilla, vanillin, creamy and phenolic	25

In addition to these classes of chemicals, the coffee volatile fraction is rich of many other compounds that are summarized in table 1.2.6

Table 1.2.6 Miscellaneous compounds characterizing the coffee volatile fraction

VOC	CAS n°	Aroma Description	Sensory Threshold (ppb)
		Aldehydes	
2-methylbutanal	96-17-3	Musty, chocolate, nutty, furfural and iaovaleraldehyde-like with malty and fermented nuances	1.3
2-methylpropanal	78-84-2	fresh aldehydic floral	-
3-methylbutanal	590-86-3	malty	0.35
(E)-2-nonenal**	18829-56-6	green, cucumber, aldehydic, fatty with a citrus nuance	0.08
3-methylpropanal	78-84-2	fresh aldehydic floral	-
acetaldehyde	75-07-0	pungent fresh aldehydic refreshing green	0.7
methylpropanal	78-84-2	fresh aldehydic floral	0.7
p-anisaldehyde**	123-11-5	Sweet, powdery, vanilla, anise, woody, coumarin and creamy with a spicy nuance	27
Phenylacetaldehyde**	122-78-1	Honey, floral rose, sweet, powdery, fermented, chocolate with a slight earthy nuance	-
propanal	123-38-6	Ethereal, pungent	10
		Ketones	
1-octen-3-one	4312-99-6	Intense earthy, metallic, mushroom-like with vegetative nuances of cabbage and broccoli. It has minor savory notes of fish and chicken	0.0036
2,3-hexadione	3848-24-6	Sweet, creamy, caramellic, buttery with a fruity jammy nuance	-
2,3-butanedione	431-03-8	strong butter sweet creamy pungent caramel	0.3
2,3-pentanedione	600-14-6	Buttery, nutty, toasted, caramellic, diacetyl and acetoin notes	20
4-(4'-hydroxyphenyl)-2-butanone	5471-51-2	Berry, sweet, woody and raspberry with a ripe, jammy, seedy character	1-10
1-(2-furanyl)-2-butanone	4208-63-3	slight rum-like	-
		Acids	
2-methylbutyric acid**	116-53-0	Acidic, fruity, dirty, cheesey with a fermented nuance	10
3-methylbutyric acid**	503-74-2	Cheese, dairy, acidic, sour, pungent, fruity, stinky, ripe fatty and fruity notes	700
		Esters	
ethyl-2-methylbutyrate	7452-79-1	Fruity, estry and berry with fresh tropical nuances	0.5
ethyl-3-methylbutyrate	108-64-5	Sweet, diffusive, estry, fruity, sharp, pineapple, apple, green and orange	0.6
		Pyridines	
pyridine	110-86-1	sickening sour putrid fishy amine	77
		Pyrroles	
1-methyl pyrrole	96-54-8	powerful smoky woody herbal	
		Terpenes	
Linalool**	78-70-6	Citrus, orange, floral, terpy, waxy and rose	0.17
limonene	5989-54-8	terpene pine herbal peppery	4
geraniol	106-24-1	Floral, sweet, rosey, fruity and citronella-like with a citrus nuance	1.1

This selection includes aldehydes, ketones, acids, esters, pyridine, pyrroles and terpenes and confirms the complexity of coffee aroma. Aldehydes comes from the Strecker degradation of amino acids and/or the autoxidation of unsatured fatty acids present in the green beans. Among the aldehydes (E)-2-nonenal, p-anisaldehyde and phenylacetaldehyde are in the list of coffee key aroma compounds and are characterized by green, fatty, sweet notes.
The ketones are produced from the pyrolysis of carbohydrates; they show quite different aromatic notes, from the mushroom like odor of 1-octen-3-one, to the caramellic sweet notes of the diketones.
Acids originate from the oxidation of the corresponding Strecker aldehydes and are responsible for fruity and fresh notes.

COFFEE FLAVOUR QUALITY EVALUATION

Defining the quality of coffee is not a simple task, several publications and coffee experts have proposed different definitions and discussions on the subject [18]
Despite this huge work, the definition has remained elusive although a rational approach to quality is becoming increasingly important also because the increasing importance of the specialty coffee community.
Cupping or cup quality would be the final determining factor for purchase or rejection of a consignment and for its price determination. The presence of defects could result in an unclean cup and thus lower the cup quality and price. [19]
Sensory evaluation comprises a set of well-established methods providing useful information about the human perception of food flavour. Cupping has always been at the heart of coffee world and production. Coffee cuppers first select green coffee and then other expert coffee tasters (the roaster, the brewer, and the barista will all do sensory assessments to judge the flavour of coffee within their own specific skills to know more about their product and its potentialities.

The key actors of coffee sensory assessments can be collected in three groups: the expert coffee tasters, the sensory analysts and the consumers. These figures are obviously involved in three different levels along the value chain and give different contributions.
Expert coffee tasters have been trained to define each single flavour note within the overall perception trying to limit the influence of the cognitive processes (e.g. memories, emotions) in their final evaluations. They provide very detailed descriptions of coffee sensory profiles that are important for coffee selection and prizing but that may seldom be perceived by the consumer.
On the other hand, the consumers, have a limited knowledge of the product but they are the final target of the whole chain and they express their evaluations in terms of enjoyment or whether or not the final product meets their expectations.
The sensory analyst role is to be the link between green coffee suppliers, roasters, and other production personnel, brewers, barista, and marketers. They are responsible for the product description that has to be done using a common language agreed upon among professionals. The accuracy of these descriptions is critical because they are the base of the product identity throughout the value chain[20]. The examination of sensory data with production parameters or consumer reference may give information of bond the technological sources of certain flavour attributes and the characteristics appreciated by the consumers.

A suitable sensory assessment is especially important in coffee because, as a natural crop product, its quality is in constant evolution, and consumer preferences evolve over time. Thanks to the continuous innovation in the discipline, methods are becoming more flexible, faster, and

adaptable to the business reality, but they require more than ever the know-how of the sensory professional(s)[20].

The Specialty Coffee Association of America (SCAA) proposed a detailed protocol to provide an accurate assessment of the coffee cup quality[21]. This cupping test, also known as "Brazilian method", is the worldwide standard method to define the quality of the incoming raw beans before the next steps of coffee transformation. It is used to select beans for specialty coffee for mono-origin or blends and to establish the price of the raw material besides the bean visual appearance properties (i.e. color, bean uniformity, and lack of "defective beans").

The cupper's quality perceptions are determined by the analysis of different specific flavour attributes that are rated on a numerical scale. The scores between samples can then be compared: coffees that receive higher scores should be noticeably better than those that receive lower scores.

The protocol reports all details about each step of the evaluation in order to maximize the level of standardization. The instructions include not only roasting and grinding parameters but also temperatures, times and number of stirring between all steps

Of particular interest is the detailed evaluation procedure reported below:

Step #1 – Fragrance/Aroma

- Within 15 minutes after samples have been ground, the dry fragrance of the samples should be evaluated by lifting the lid and sniffing the dry grounds.
- After infusing with water, the crust is left unbroken for at least 3 minutes but not more than 5 minutes. Breaking of the crust is done by stirring 3 times, then allowing the foam to run down the back of the spoon while gently sniffing. The Fragrance/Aroma score is then marked on the basis of dry and wet evaluation.

Step #2 – Flavour, Aftertaste, Acidity, Body, and Balance

- When the sample has cooled to 160º F (71º C), in about 8-10 minutes from infusion, evaluation of the liquor should begin. The liquor is aspirated into the mouth in such a way as to cover as much area as possible, especially the tongue and upper palate. Because the retro nasal vapors are at their maximum intensity at these elevated temperatures. Flavour and Aftertaste are rated at this point.
- As the coffee continues to cool (160º F (71.1°C) - 140º F (60°C)), the Acidity, Body and Balance are rated next. Balance is the cupper's assessment of how well the Flavour, Aftertaste, Acidity, and Body fit together in a synergistic combination.
- The cupper's preference for the different attributes is evaluated at several different temperatures (2 or 3 times) as the sample cools. To rate the sample on the 16-point scale, circle the appropriate tick-mark on the cupping form. If a change is made (i.e. if a sample gains or loses some of its perceived quality due to temperature changes), re-mark the horizontal scale and draw an arrow to indicate the direction of the final score.

Step #3 – Sweetness, Uniformity, and Cleanliness

- As the brew approaches room temperature (below 100º F (37.8°C)) Sweetness, Uniformity, and Clean Cup are evaluated. For these attributes, the cupper makes a judgment on each individual cup, awarding 2 points per cup per attribute (10 points maximum score).
- Evaluation of the liquor should cease when the sample reaches 70º F (21º C) and the *overall score* is determined by the cupper and given to the sample as "Cupper's Points" based on ALL of the combined attributes.

Step #4 – Scoring

After evaluating the samples, all the scores are added as described in the "Scoring" section below

and the Final Score is written in the upper right hand box

Although sensory evaluation is normally carried out by trained panelists, there are some limitations such as low repeatability and reproducibility in the results because of many subjective (e.g., sensory susceptibility of the person, health, fatigue) and objective (conditions under which the analysis is performed) parameters, as well as the time needed for its implementation.

In this perspective, several studies have been focused in the investigation of the relationship between sensory and compositional properties of food matrices and in particular about coffee.
This thesis deals with a multidisciplinary study on the relationship between the chemicals of the coffee flavour and its sensory properties. The knowledge of the chemical signature of different coffee aroma notes linked to their sensory perception to define the quality of the beans can be used as an objective tool in routine control in view of supporting the panel in cupping routine work in the sensory assessment of the coffee quality.

References:

1. Spence, C. Comfort food: A review. *Int. J. Gastron. Food Sci.* **9,** 105–109 (2017).
2. Otake, K. & Kato, K. Subjective Happiness and Emotional Responsiveness to Food Stimuli. *J. Happiness Stud.* **18,** 691–708 (2017).
3. Locher, J. L., Yoels, W. C., Maurer, D. & van Ells, J. Comfort Foods: An Exploratory Journey Into The Social and Emotional Significance of Food. *Food Foodways* **13,** 273–297 (2005).
4. Cizza, G. & Rother, K. I. Was Feuerbach right: are we what we eat? *J. Clin. Invest.* **121,** 2969–2971 (2011).
5. Buettner, A. *Handbook of Odor*. (Springer, 2017).
6. Sunarharum, W. B., Williams, D. J. & Smyth, H. E. Complexity of coffee flavor: A compositional and sensory perspective. *Food Res. Int.* **62,** 315–325 (2014).
7. Ribeiro, J. S., Ferreira, M. M. C. & Salva, T. J. G. Chemometric models for the quantitative descriptive sensory analysis of Arabica coffee beverages using near infrared spectroscopy. *Talanta* **83,** 1352–1358 (2011).
8. Lindinger, C. *et al.* When Machine Tastes Coffee: Instrumental Approach To Predict the Sensory Profile of Espresso Coffee. *Anal. Chem.* **80,** 1574–1581 (2008).
9. Caprioli, G., Cortese, M., Sagratini, G. & Vittori, S. The influence of different types of preparation (espresso and brew) on coffee aroma and main bioactive constituents. *Int. J. Food Sci. Nutr.* **66,** 505–513 (2015).
10. Caprioli, G. *et al.* Optimization of espresso machine parameters through the analysis of coffee odorants by HS-SPME–GC/MS. *Food Chem.* **135,** 1127–1133 (2012).
11. Belitz. Hasn-Dieter, Grosch, W. & Schieberle, P. *Food Chemistry*. **53,** (Springer, 2013).
12. Oestreich-Janzen, S. Chemistry of Coffee. in *Comprehensive Natural Products II* 1085–1117 (Elsevier, 2010).
13. Wang, H. Y., Qian, H. & Yao, W. R. Melanoidins produced by the Maillard reaction: Structure and biological activity. *Food Chem.* **128,** 573–584 (2011).
14. Frank, O., Zehentbauer, G. & Hofmann, T. Bioresponse-guided decomposition of roast coffee beverage and identification of key bitter taste compounds. *Eur. Food Res. Technol.* **222,** 492–508 (2006).
15. Grosch, W. Flavour of coffee. A review. *Nahrung / Food* **42,** 344–350 (1998).
16. Audouze, K. *et al.* Identification of Odorant-Receptor Interactions by Global Mapping of the Human Odorome. *PLoS One* **9,** e93037 (2014).
17. Flament, I. *Coffee Flavor Chemistry*. (Wiley 2001).
18. Illy, A. & Viani, R. *Espresso Coffee-The Science of Quality*. (Elsevier Academic Press, 2014).
19. International Trade Centre. *The Coffee Exporter ' s Guide*. (2011).
20. Britta Folmer. The Craft and Science of Coffee. in *The Craft and Science of Coffee* (eds. Blank, I., Farah, A., Giuliano, P., Sanders, D. & Wille, C.) iii (Elsevier, 2017).
21. SCAA,Protocols Cupping Specialty Coffee. *SCAA Protoc.* 7 (2008).

1.3 The role of analytical chemistry in food analyses: platforms for the analysis of food flavour in foodomic studies

Food analysis is, nowadays, one of the most important application field of analytical chemistry due to the even-increasing need to be in compliance with legal standards on health risks, environment, animal welfare and ethical requirements. In addition, food analyses are becoming crucial in the assessment of food sensory impact that is becoming even more a key parameter for the quality evaluation.

Food quality is an extended concept that is traditionally mainly related to food safety. For this reason, the first goal of food analysis has been and still is to ensure this essential criterion.
In addition to food safety, food quality definition pass through the evaluation of several other characteristics for which analytical chemistry plays a crucial role. For instance, i) food analyses are important in detecting adulteration and frauds; ii) in the analysis of microbiological properties, iii) in the evaluation of food (or part of it) biological effects iv) and in the study of food rheology, morphology, structure or surface and its chemical composition is fundamental to standardize a product or to improve its characteristics or, else, to design new foods.

The definition of food quality is undoubtedly becoming a challenging and complex issue because in recent years it has moved from the "simple" concept of healthy to a more holistic food characterization. The development of new analytical strategies (informative, fast, automated, and sensitive) and applications, has grown together with the knowledge the consumers have acquired about what they eat or drink.
The knowledge of the chemical composition of a food is therefore essential to understand the changes occurring along the whole production chain, from the raw materials up to the final-products and to evaluate their shelf-life.

Another important challenge in food analysis is to improve our limited understanding of the roles of food compounds at the molecular level (i.e., their interaction with genes and their subsequent effect on proteins and metabolites) for the rational design of strategies to manipulate cell functions through diet, which is expected to have an extraordinary impact on our health and well-being.
Food scientists are therefore pushed to develop new tools, standards and approaches also moving from their classical procedures to modern analytical techniques that allowing them to give an adequate answer to the global quality control demand.

As in depth discussed in the previous sections a great portion of food quality definition involves its sensory impact summarized by the concept of *flavour* (more details in section 1.1).
The key role of chemical senses as smell and taste in food flavour definition and in food decision making procedures confirms the importance of the food chemical composition in this complex scenario[1]. Because of its complexity, food flavour chemical characterization cannot be based only on a list of preselected compounds. Moreover, its conventional definition considers the features

under study singularly without taking into account the relationships among them and how characteristics influence one another.
The efforts in this direction lead to the development of analytical approaches, defined as –omic approaches, able to achieve a global food characterization and to define its chemical, physical and biological characteristics.
Thanks to the omics approach, it is now possible to connect food components, diets, individuals, health, and diseases, but this broad vision needs not only the application of advanced technologies, but also and mainly the ability of looking at the problem with a different approach, a ''foodomics approach''[1,2].

"Foodomics" is a comprehensive strategy that has many different definitions according to the field of application; someone says that foodomics is "a discipline that studies the food and nutrition domains through the application and integration of advanced omics technologies to improve consumer's well-being, health and confidence", others define it as "the comprehensive, high-throughput approach for the exploitation of food science in the light of an improvement of human nutrition" or, else, as "a new approach to food and nutrition that studies the food domain as a whole with the nutrition domain to reach the main objective, the optimization of human health and well-being".[3] Despite these definitions differs more or less markedly one another, all of them share a comprehensive multidisciplinary characterization of food that will give information and tools useful to improve the consumers health and well-being. According to the –omic approach this characterization requires the integration of advanced analytical techniques and bioinformatics. Compared to the past, the Foodomic approach requires a very important data elaboration step that affords to transform the information and the data collected from foods into results.

The number of opportunities (e.g., new methodologies, new generated knowledge, new products, etc.) derived from this trend are impressive and for example, it includes the possibility to account for food products tailored to promote the health and well-being of groups of population identified on the basis of their individual genomes. The introduction, in this area of research, of advanced "omics" approaches such as Foodomics have allowed food scientists to face problems unthinkable a few years ago.
Just to make some examples, the –omic approach allows:

- to understand the effect of bioactive food constituents on crucial molecular pathways and understand the bio-chemical, molecular and cellular mechanisms behind the beneficial and adverse effects of these components,
- to determine the genetic differences among individuals in response to specific nutrients or sensory stimuli and to understand how the diet can be modulated in the perspective of diseases prevention or well-being improvement,
- to understand the molecular basis of biological processes with agronomic interest and economic relevance, such as the changes during food shelf-life, food consumption and the stress adaptation responses of/to food pathogens to improve food hygiene, processing and conservation.
- the comprehensive assessment of food safety, quality and traceability.

In spite of the huge potential of foodomics, it has to be highlighted the methodological difficulties to carry it out. The approach is not easy and requires a high degree of complementary knowledges coming from different fields, typically including analytical chemistry, biology/medicine, bioinformatics, and statistics.

ANALYTICAL PLATFORMS FOR FOOD FLAVOUR IN FOODOMIC EVALUATIONS

The approach to analysis has radically changed with the introduction of -omic sciences. The development of global and high-throughput analytical strategies has been supported by the rapid evolution of analysis and detection techniques managing high numbers of samples and screening simultaneously several analytes.
Moreover, these needs prompted the development of new approaches for sample preparation and analysis and led to the definitive inclusion of data elaboration as an active step of the entire analytical procedure.
Foodomic analyses provide highly complex data sets that requires heavy treatments to be translated into results [4,5].
Both routine and research laboratories, where large numbers of homogeneous samples are processed every day, have therefore focused their investigations on the development of automatic systems, in which "sample preparation, analysis data and interpretation" are merged in a single step, in the so-called total analysis systems (TASs)[6].
During the development of an analytical protocol each block must be chosen according to i) the aim of the project, ii) the matrix investigated and iii) the required level of information.

Sample Preparation: a key step of food analyses

Over the last years, several and important results have been achieved in instrumentation development achieving ever better performance.
The same effort has been made for sample preparation. Sampling is the first step of the analytical process and it is crucial for the correct interpretation of the results; an error at this stage can no longer be corrected and affects the whole analytical process[7,8,9].
Advance in sample preparation should minimize solvent use and hazardous waste production, save time and cost per sample, and, at the same time, improve the efficiency of the analytes isolation.

Foods are very complex samples and require specific sample preparation procedures able to exploit i) the chemical and physical properties of the compounds of interest (volatility, polarity, solubility and thermal, oxidative, and hydrolytic stability), and ii) the interactions of the analytes with other components in the sample (matrix).
A rapid analyte isolation is essential to minimize or prevent the typical alterations occurring in food samples because of enzymatic activity, lipid oxidation, microbial growth and physical changes.
Sampling of liquid matrices must ensure their uniformity and use non-invasive methods (e.g., avoid the formation of precipitates by mild heating).

The most critical aspect of sample preparation in flavour analyses concerns the volatile compounds that is typically based on their volatility and/or solubility. The sampling techniques suitable for -omic studies have to fulfil some key requisites:

- provide a consistent picture of the informative analytes
- minimize the artefacts formation,
- tune the extraction parameters by optimizing physico-chemical operative conditions (temperature, time, pressure etc..) ,
- be reproducible and effective enough to match the instrumental sensitivity,
- be simple, fully automatable and suitable to be integrated in the analytical system,
- be environmental and economically sustainable

Traditionally, the extraction of volatiles have been performed by extraction/concentration techniques like simultaneous distillation/extraction (SDE)[10], steam distillation techniques, direct organic solvent extraction, solvent assisted flavour evaporation (SAFE)[11] or even with ultrasounds and microwaves assisted techniques.
The extraction conditions have to be optimized to maximize the speed and the recovery of the analytes. If necessary, the extraction step can be followed by a purification step to remove the inferring molecule or reduce the matrix complexity.
In spite of their effectiveness, in the flavour field, these techniques are laborious, not environmental friendly, and, above all not suitable for an "–omic" food flavour analysis.

A great development in this field has been represented by the solid phase extraction (SPE); this extraction is based on the use of specific cartridges or columns pre-packed with various stationary phases (reverse phase C4, C8, or C18, ion exchange etc).
The liquid sample is loaded through the SPE column where target analytes are selectively retained from the sorbent while other interfering and matrix components are washed out with dedicated solvents or vice versa.
Miniaturized SPE, specifically designed for sample preparation in foodomic analysis, are commercialized by several manufacturers and have become increasingly popular because of their easy and fast use. They allow single-step desalting, concentration, and purification of food samples for sensitive downstream analyses, for instance proteomics and metabolomics.
Solid Phase Micro-extraction (SPME) technique is an example of solid phase extraction largely used in food and beverages analyses.
SPME belongs to the High Concentration Capacity (HCC) techniques.
These techniques were introduced by Arthur and Pawliszyn in the early 90's. The idea behind them is to keep the advantages of SPE while increasing the technique's sensitivity (i.e. increasing the concentration factor CF).
In the HCC techniques, volatiles and semi-volatiles are accumulated and extracted from the matrix (liquid or gaseous) into a polymeric coating of different nature and then recovered by liquid or thermal desorption. In addition, the HCC techniques often avoid the use of solvents making them environmentally friendly.
The HCC techniques exploit two main extraction mechanisms (depending on the nature of the polymer) to capture the analytes: sorption or adsorption[9].
During sorption, the analytes are partitioned in the polymeric material while during adsorption the analytes are retained on the polymers active surface where a fixed number of adsorptive sites are present. In adsorption, a competition between the analytes can occur, since the number of bonding sites is limited.
The selection of the polymer not only conditions the extraction principle but also should be done according to the chemical properties of the analytes.
Polydimethylsiloxane (PDMS) is one of the most common polymer coating used in HCC techniques; it acts by sorption and it is most suitable for non-polar analytes.
Polyacrilate (PA) and Divinylbenzene (CVB) are used to recover polar and highly polar analytes.
The combination of polymers operating in different modes have been applied to several HCC devices because it extends the range of polarity and/or increases the analyte recovery by exploiting both the adsorption and sorption mechanisms.
The HCC techniques have been chosen as sampling approach in the experimental part of this thesis. In particular, Solid Phase Microextraction and the Stir Bar Sorptive Extraction will briefly be described below.

Direct Immersion-Solid Phase Microextraction (di-SPME)

SPME was the first HCC sampling techniques introduced by Pawliszyn in 1990; it is a solvent-free sampling technique where the analytes are sorbed and/or adsorbed exposing the sampling material to the sample. The device consists of a polymer-coated fused silica fibre, which is assembled in a dedicated needle of a special holder (Figure 1.3.1).

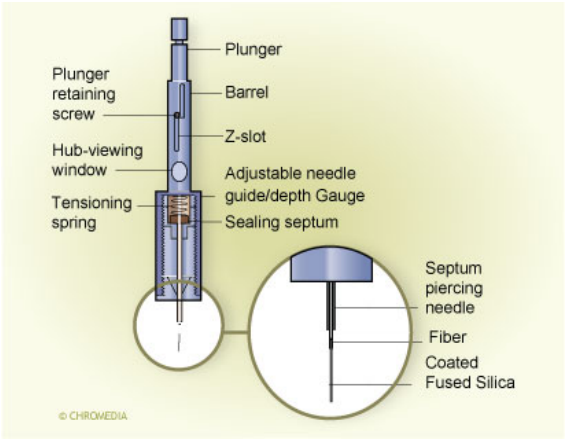


Figure 1.3.1 Scheme of the conventional SMPE device (www.chromedia.org)

During sampling (Extraction) the fibre is immersed in the liquid sample by pushing the plunger out of the device. Analytes are then extracted from the matrix and concentrate into the polymer coating and then transferred in the analytical system by thermal desorption at high temperatures or liquid desorption by solvents (Figure 1.3.2).

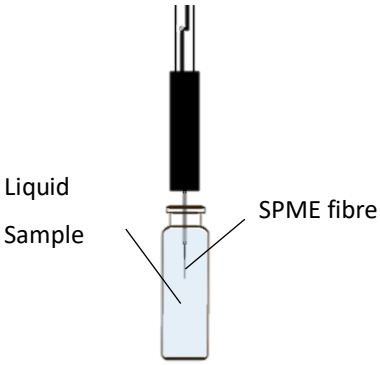


Figure 1.3.2 Direct immersion-SPME sampling scheme

One of the main advantages of the SPME is the possibility to choose the polymeric coating depending on the analytes under study.
Fibres can be coated with a single material such as polydimethylsiloxane (PDMS) and polyacrylate (PA) or with multi components materials, in particular PDMS (e.g Polydimethylsiloxane/Divinylbenzene/Carboxen DVB/CAR/PDMS) to extend the range of recovered analytes (in terms of polarity and volatility) and to exploit both sorption and adsorption mechanisms.

Thanks to its simplicity and versatility, the use of SPME in food analysis increased exponentially, in particular in the studies on food aroma composition and contaminants; this flexibility is due to the

different and new polymeric coatings affording to cover a wide range of polarity and volatility as well as when designed for specific applications (i.e. – polymeric ionic liquids)[12].
In spite of its diffusion, SPME show a relatively limited concentration capacity because of the small volume of polymers involved in the extraction.

Solid Phase Sorptive Extraction (SBSE)

This technique has been developed in the 1999 to overcome the main limit of the SPME: the concentration capacity.
SBSE is based on a glass-coated magnetic stir bar in its turn coated by a thick film of polymers in which analytes are concentrated. The device is commercially known as Twister®. Sampling is done by directly introducing the SBSE device into the aqueous sample; in the original experiments, the analytes sampled during a given time were recovered by thermal desorption and then on-line transferred to a gas chromatographic (GC) or a GC–MS (mass spectrometry) system for analysis. Later, liquid desorption in combination with high performance liquid chromatography (HPLC) was also applied, mainly for analytes not compatible with GC. (Figure 1.3.3)
The most common polymeric coating is PDMS that affords to recover non-polar and medium to high volatility compounds at high rate. The PDMS apolarity is its main limit that can be overcome by combining two or more materials operating on different principles (e.g., sorption and adsorption) to capture a wider range of compounds (e.g: Polydimethylsiloxane (PDMS)/polyethylene glycol (EG); Polydimethylsiloxane (PDMS)/Polyacrilate (PA) etc.) [9,13,14].
Another approach includes the so-called dual-phase (DP) Twister that consists of a PDMS tube with an inner cavity packed with adsorbent materials such as Carbopack B, Tenax GC, a bisphenol-PDMS copolymer and Carbopack coated with 5% of Carbowax [15]. The concentration capability of dual-phase twisters is therefore the result of the sorption of the analytes onto PDMS from liquid or vapour phase, followed by their diffusion through the PDMS layer and the adsorption onto the inner phase. The effectiveness of dual-phase stir bars therefore depends on the permeability of the outer PDMS layer to the investigated analytes, the adsorption capability of the inner material, and, last but not least, the strength of the analyte–inner phase interaction, that must be reversible to afford their total release through thermal desorption[14, 21, 16].

Compared to di-SPME, SBSE provides a better concentration capability thanks to the higher amount of polymer coating the twister but, on the other hand, its main limited is the small number of polymer coatings commercially available, the analyte recovery from the polymers after liquid sampling and the need of specific and expensive devices for stir bar thermal desorption (TDU or TDS).

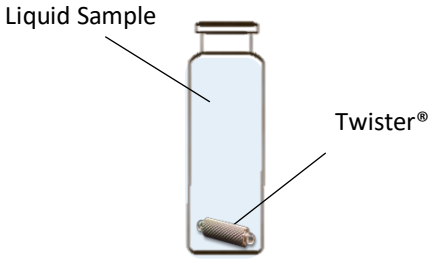


Figure 1.3.3 SBSE sampling scheme

High Concentration Capacity (HCC) sampling techniques applied to sample headspace (HS)

Before discussing applications and advantages of HCC techniques the basic concept of headspace sampling and the physical laws that rule its composition have to be described.
Headspace (HS) Sampling techniques exploit a peculiar property of aroma compounds i.e. their volatility[17].
Considering a sealed vial, the headspace refers to the vapour phase directly above the sample (solid or liquid). The idea behind Static Headspace sampling is to reach the equilibrium between the condensed and the vapour phases; therefore, in many cases the need of not too long analyses (suitable for routine purposes) has to be met. Non-equilibrium conditions are by definition exploited in dynamic Headspace sampling (D-HS) where the volatile compounds are continuously extracted from the matrix using a gas as carrier[18].

$$A(vap) \leftrightarrow A\ (cond) \text{ [1]}$$

$$K = \frac{[A(cond)]}{[A(vap)]} \quad \text{[2]}$$

The equilibrium between vapour and condensed phases is summarized in equation [1] where A is the analyte of interest.
The analyte partition between the vapour and the condensed phase is ruled by the partition coefficient (K) defined by equation [2].
Another important factor affecting the partition of the analytes in the headspace is the phase ratio (β) defined as the ratio between the volumes of the gaseous and the condensed phases [3]:

$$\beta = \frac{V(vap)}{V(cond)} \text{ [3]}$$

At the equilibrium, K and β are constant and the composition of the gas phase is representative of that of the condensed phase (sample).
The partition coefficient K depends on several variables:

- the temperature,
- the analyte vapour pressure,
- the activity coefficient of the analyte (*matrix effect*).

These parameters are ruled by three basic laws: *Dalton's*, *Raoult's* and *Henry's* laws.

Dalton's law states that the total pressure of as mixture of gases (the headspace of every matrix) is the sum of the partial pressures of i^{th} components of the mixture:

$$p_{total} = \sum_{i=1}^{n} p_i \text{ [4]}$$

the partial pressure of each analyte can be related to its molar fraction (x_i) through the Raoult's law:

$$p_i = kx_i \quad where\ k = p_i^o \text{ [5]}$$

the proportion constant (k) is function of p_i^o; the vapour pressure of the analyte at a given temperature.
The proportion constant k is usually known as Henry's constant (H).

These rules therefore are valid just for ideal solutions and not for non-ideal solutions that are the majority of real-world samples.
To compensate this equation the activity coefficient (γ_i) has been introduced.
This coefficient considers the real portion of the molar fraction of an analyte really available in the vapour phase. This factor depends on the nature of the analyte and reflects the intermolecular interactions between the analyte and other sample components, for example the solvent or the texture or the composition of the matrix.

$$p_i = p_i^o \gamma_i x_i \text{ [6]}$$

In other words, the partial pressure of an analyte in the headspace (p_i) is given by the interaction between its vapour pressure (p_i^o), the activity coefficient of the matrix (γ_i) and its molar fraction (x_i).
Furthermore, the vapour pressure (p_i^o) and the activity coefficient (γ_i) are inversely proportional to the partition coefficient (K) that rules the distribution of the analyte between the condensed and the gaseous phases.
The vapour pressure is linked to the temperature as shown in equation:

$$log p_i^o \propto \frac{B}{T} + C \text{ [7]}$$

where B and C are specific constants of the analyte and T is the temperature. This equation suggests that the increase of the sampling temperature implies the increase the vapour pressure and thereby of the headspace sensitivity.

In summary, the headspace sensitivity (α) is strictly related to the phase ratio (B) and the partition coefficient (K) as confirmed by the equation:

$$\alpha = \frac{1}{K+B} \text{ [8]}$$

The parameters to tune to increase the headspace sensitivity are therefore:

- the temperature of the system;
- the volume of the sample (i.e. the phase ratio);
- the chemical characteristics of the matrix (e.g. salting out effect);
- the chemical properties of the analyte (changing its volatility through derivatization).

Since its introduction, static headspace sampling (S-HS) in combination with GC has successfully been employed in the flavour and fragrance field for several applications.
Its success was due to its simplicity, rapidity and to the possibility of automation that this sampling approach offers. Despite these important advantages, the S-HS main limitation is its low sensitivity. This characteristic limits its application between ppm and percent concentrations.

The introduction of high concentration capacity techniques (HCC) has increased the complexity of the system, since the sampling material influences the recovery equilibrium. In HCC techniques two equilibria have to be considered: 1) the partition of the analytes between the matrix (liquid or solid) and the vapour phase, 2) the partition of the analytes between the vapour and the polymeric phases.

Consequently, further parameters, in addition to those used in the headspace, have to be considered[19]:

- the partition coefficient between the headspace and the polymer,
- the nature of the polymer, its volume and its size,
- ionic strength.

One of the most popular HCC-HS sampling technique is the headspace-Solid Phase Microextraction (HS-SPME). This technique found many applications in food flavour analyses because it exploits the concentration capacity of SPME and therefore enhances the sensibility of conventional S-HS.
The quantitation accuracy and reliability of HS-SPME technique has been investigated and demonstrated in 2011 by Bicchi et al.[20]. Furan in roasted coffee was analysed by HS-SPME-GC-MS and quantified by Multiple Headspace Extraction (MHE), Standard Addition (SA) and Stable Isotopes Dilution Analysis (SIDA) showing consistent results.

To improve the sensitivity of HS-SPME due to the low volume of fiber coating the use of the SBSE was extended almost immediately to vapour phase sampling (headspace) by Bicchi et al.[21] and Tienpont et al.[14] with the name of Headspace Sorptive Extraction (HSSE). In HSSE, sampling is in static mode by suspending the PDMS stir bar in the vapour phase in equilibrium (or not) with the solid or liquid matrix. Again, accumulated analytes are recovered by thermal desorption combined and on line analysed by GC or GC–MS.
A schematic view of HS-SPME and HSSE sampling techniques is reported in figure 1.3.4.

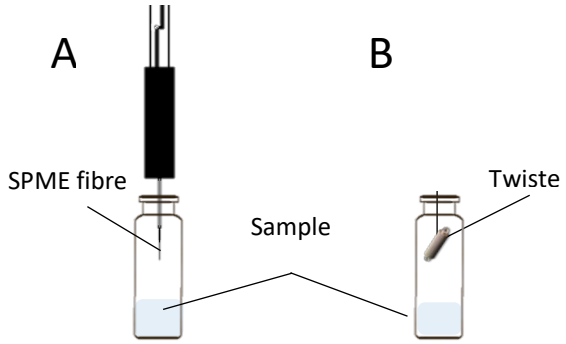


Figure 1.3.4 Schematic view of HS-SPME (A) and HSSE (B) sampling techniques

Analytical platforms for food analyses

Core of each TAS is the analytical platform. In the recent decades, the technological development made available new and powerful platforms providing ever more extended information.
Analysis is addressed to study several aspects of a food therefore requiring different analytical platforms including (i) spectroscopic techniques such as mass spectroscopy (MS), nuclear magnetic resonance (NMR), infrared (IR), etc.; (ii) separation techniques such as high performances liquid chromatography (HPLC) and gas chromatography (GC), capillary electrophoresis (CE), supercritical fluid chromatography (SFC) mainly combined with mass spectrometry, (iii) biological, immunological techniques, etc.
Despite the spectroscopic techniques are gaining importance and are widely used in particular in quality control, the separation techniques still are still the most valid and effective tool for the –omic food analysis[22,23].
The development of spectroscopic techniques can be explained because they can be easily coupled with chemometric and statistic data treatment. These data elaboration tools are appropriate and useful for the evaluation of the encrypted information deriving from the spectroscopic analyses.
Separative techniques still represent a reference for the analysis of foods since they afford to separate the components in complex mixtures and the identification of the pool of chemicals responsible for the characteristics under study.
The choice of the analytical platform has to be done in agreement with the level of information needed; today the scenario is wide and can be splitted in two main classes: chromatographic and direct spectroscopic techniques.

Chromatographic techniques & detection

The advent of modern chromatography in the 50's[23] changed food analysis affording the separation of the components of complex matrices.
Food analyses mainly exploit gas (GC) and liquid chromatography (LC).
The introduction of GC allowed the separation, identification and determination of volatile or volatilizable chemical compounds in complex mixtures. On the other hand, the potential of LC, which can differently be approached (e.g. reverse phase, ion exchange, affinity, etc.), has been enormously increased with the introduction of HPLC (early 70') and more recently (late 90') by ultra-high-performance liquid chromatography (UPLC).
Chromatographic techniques can be associated with a large array of detectors based on different principles and able to detect a wide range of compounds with different chemical structure or to be selective for specific structural characteristics[3].
The data acquired from a chromatographic method include retention times, peak intensity (area or height) and mass spectra, of course when in combination with MS.
When combined with chromatography, the MS fragmentation differs depending on the interface and ionization mode, and on the architecture of the analyser (high or low-resolution MS).
In GC, where direct electron impact (EI) is the most common, the fragmentation provides a great number of more or less diagnostic fragments, whose pattern can be used to identify the analytes by comparison with those contained in commercial or home-made libraries.
In LC the fragmentation depends on interface and ion source and the resulting MS pattern has a different diagnostic impact; generally, LC-MS is characterized by a less fragmentation that can be exploited in isotopic peaks detection[24].

The versatility of the chromatographic methods affords the analyst to interact on both separation and detection to acquire an analytical set of data with at the maximum level of information, as well as to select the data suitable for a further treatment.

Gas Chromatography (GC-FID, GC-O, GC-Ms)

In gas-chromatography, the analytes must be vaporizable (or derivatized to make them volatile) and thermo-stable since temperature conditions injection and the chromatographic process).
The stationary phase is bonded or coated on the inner surface of a fused silica capillary column. In gas-liquid chromatography (GLC), the most widely used coating materials are highly-viscous non-volatile polymeric liquids unlike in gas solid chromatography (GSC) where it is solid:
In a GC system, the sample is introduced in the column through an injection system; the column is placed into a temperature-controlled oven where the separation take place: the end of the column is connected to a detector.
During the chromatographic run, each analyte is eluted in agreement with its distribution coefficient that drives the partition between the stationary and the mobile phases. During the run the molecule are move forward by the carrier gas.
One the most important GC advantages is its flexibility; it can be coupled to several detectors with different characteristics.

GC-FID: Flame Ionization Detector is one of the simplest and universal detectors used in GC: the signal originates from the ions generated during the combustion of organic compounds in a hydrogen-rich flame that are submitted to a polarizing voltage generating a current proportional to the amount of eluting analyte.
A scheme of the FID detector is reported in figure 1.3.5:

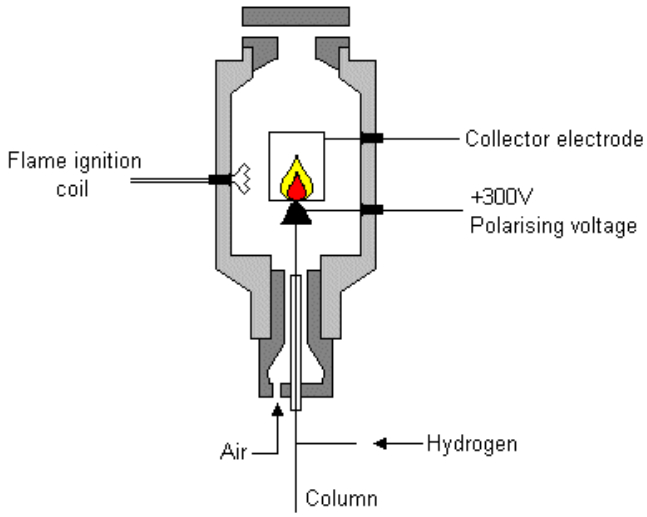


Figure 1.3.5 Scheme of a FID "www. teaching.shu.ac.uk"

Because of its simplicity the GC coupled to a FID found several applications in food analyses, in particular when (i) high sensitivity, reproducibility, and stability, (ii) inexpensive and (iii) easy, rapid instrumentation is needed.
Over the last ten years the methods of the –omic approach has been adopted in food quality determination of different food matrices using FID as detector. Thanks to its routine suitability, a large amount of work has been carried out to relate the chemical profiles to the sensory data for predictive purposes[25,26]. The chemical fingerprints by GC-FID were widely used to investigate for example the study the ageing of wine, or to classify beers according to a portion of their volatile

fraction. Other examples involved beer, wine, cheese, green tea, balsamic vinegars and coffee[25,26,27,28,29,30].

All these papers highlight the FID advantages and demonstrate its usefulness in particular in food production sites; reasons behind this statement are the following:

- GC-FID is inexpensive and therefore available even in small factories;
- FID is a universal detector, sensitive for a wide variety of organic compounds;
- FID has a wide dynamic range and affords analyte quantitation over a wide range of concentrations;
- FID provides accurate quantitative data since its range of acquisition frequency is very wide unlike many routine MS detectors(quadrupoles).

Despite these practical advantages FID provides only limited information on the chemical nature of the analytes effectively separated.

All studies previously reported indeed combine the GC-FID fingerprinting to a GC-MS investigation of the peaks in order to identify target compounds.

A further dimension of data is sometimes added to the FID in the platform: Gas Chromatography-Olfactometry (GC-O) combines the GC separation to the detection of the human nose. In GC-O trained analysts (sniffers) detect, measure the intensity and provide a sensory description of each peak eluting from the GC column.

In GC-O, the gas flow eluting from the chromatographic column is splitted in two, the first lead to the sniffing port while the second is connected to a FID or a MS detector in order to obtain a reference instrumental trace.

A schematic view of the instrumentation is reported in figure 1.3.6.

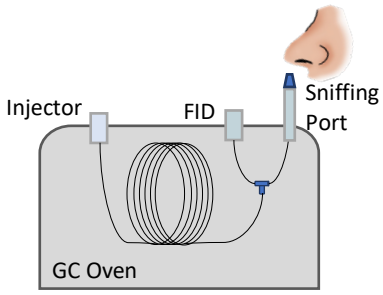


Figure 1.3.6 Schematic view of a GC-O/FID instrumentation

Olfactometry is used for screening purposes. Two main screening techniques have been developed over the years: Aroma Extract Dilution Analysis (AEDA) and Comprehensive High-throughput Array for Relative Methylation (CHARM) analysis developed by *Grosh*'s et al and *Acree*'s groups respectively. These techniques both evaluate a series of rationally diluted samples dilution of an original aroma extract using GC/O to select the characterizing odorants in it. AEDA will more in depth be discussed in a dedicate session (2.3) in the experimental part.

A more quantitative data analysis have been obtained by GC-O with Gas Chromatography- Surface of Nasal Impact Frequency (GC-SNIFF) developed by *Chaintreau* et al. in 1997[31]. In this method, the potencies of the odorants are based on the frequencies of detection of the odorants perceived at the sniffing port.

The GC-O is not properly an –omic friendly technique since needs long analyst's training and does not afford the screening of a large number of samples in short periods of time. However, its application is important to define the odour description or potency of particular odorants of off-flavour arising from an –omic study.

GC-MS: mass spectrometry, in combination with separation techniques such as chromatography and with statistical and bioinformatic tools, is today a key analytical technique on which the –omic approaches are based[3]. Mass spectrometric detection is a universal detector that usually improves the selectivity of an analytical system, because it extends the GC ability to study a complex mixture by providing a further and orthogonal information specific for each separated analyte (i.e., its spectrum). Mass spectrometers are based on the ionization and fragmentation of a compound (for examples through electron impact (EI)), followed by the physical separation and detection of the charged diagnostic ions (fragments).
A mass spectrometer mainly consists of an ion source (site of the ions generation), an analyzer (where fragments generated in the ion source are separated according to their m/z ratios) and a detector. MS data are recorded in a 2D plot displaying the ion intensity vs. the mass-over-charge (m/z) ratio known as mass spectra.
Mass spectra are highly indicative of the analyte original structure, being virtually the diagnostic fingerprint of a molecule. The method is thereby a powerful tool to identify sample components.
The most popular Mass Analyzer used for food analyses is the Quadrupole Mass Spectrometer (qMS). It consists of four cylindrical rods, set parallel to each other. The disposition of these rods creates a tunnel in which the ions produced in the ion source are separated according to their mass-to-charge (m/z) ratio. Oscillating electric fields and constant radiofrequency are alternatively applied to the rods in alternative combination and only the ion with the m/z ratio compatible with the electric field will be able to pass through the quadrupole and reach the detector.
MS quadrupoles are able to cover a mass range between 2 and 4000 amu; acquire data in Full-scan and/or in SIM or MIM (Single or Multiple Ion Monitoring) modes and provide a relatively fast scan.
qMS can be successfully applied to both conventional and multidimensional GC thanks to their ability to achieve a 20000 amu/s scan speed and 50Hz scan frequency using a restricted 290amu mass range [7,32] but sufficient for highly volatile compounds.
Figure 1.3.7 report a scheme of a MS Quadrupole analyzer

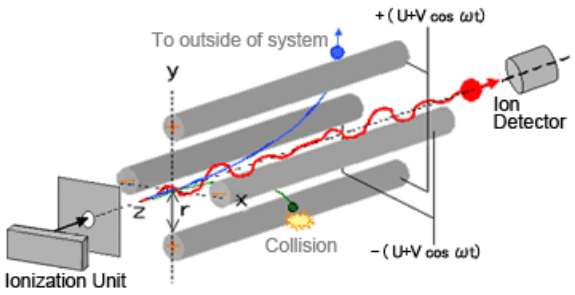


Figure 1.3.7 MS Quadrupole analyser "www.Schimadzu.com"

The MS quadrupole analyzed results to be a flexible, efficient and relatively cheap analyzer and therefore found application in many foodomic studies.
Over the last years the qMS have been applied to many studies aiming to classify, predict, identify and authenticate different food matrices.
The development of Time of Flight MS (TOF-MS) analyzers has extended the possibilities MS as detector for GC to high-speed or high-resolution GC-MS.
The most flexible geometry is the orthogonal acceleration time of flight mass spectrometer (oa-TOF MS); this analyzer can successfully be coupled with several ion sources including the electron and chemical ionizations (EI and CI).

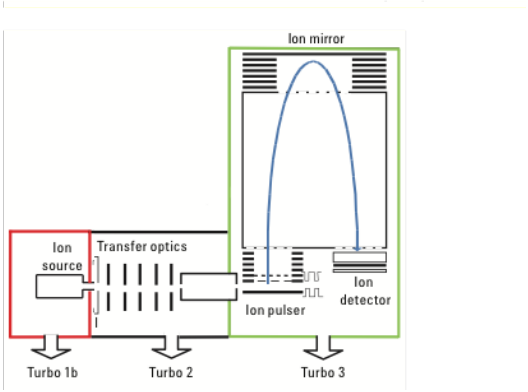


Figure 1.3.8 Scheme of an Agilent TOF MS analysed www.agilent.com/chem

Figure 1.3.8 reports a scheme of a oa-TOF MS analyser. In this instrumentation, ions generated in the ion source are shaped in a parallel beam by the transfer ion optics and directed to the fight tube where mass separation take place.
A TOF-MS analyser discriminates the ion fragments on their flight time. The ion beam reaches the ion pulser, a stack of plates each (except the back plate) with a centre hole, where, the ions are accelerated in the flight tube by a high voltage pulse. The ions travel through the flight tube (1m long), are reflected by a two-stage, electrostatic ion mirror and reach the detector where their m/z ratio is measured.

When dealing with complex sample matrices, such as food, adequate mass resolution is often essential.
High resolution mass spectrometry can be performed by a q-TOF instrumentation (Figure 1.3.9).
Unlike the conventional TOF-ms the ion source is followed by a mass filter (generally a quadrupole) where the ions are separated and sent in the collision cell.
The collision cell consists of six small parallel metal rods to witch a Radio Frequency (RF) is applied to confine the ions with a particular mass in the open center of the rod set.
A collision gas (generally Argon) in the cell enables collision induced dissociation (CID) and forms new ions from those selected by the radio frequency.
At the exit of the collision cell, ions are shaped in a parallel beam by the transfer ion optics and directed to the flight tube to be separated.

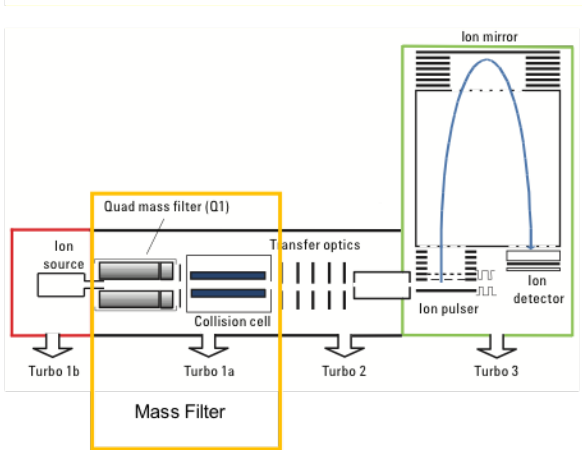


Figure 1.3.9 Scheme of an Agilent q-TOF MS analysed www.agilent.com/chem

Two types of GC–TOFMS are currently available[33]:

1. GC coupled to high-resolution (HR)TOF MS achieving mass accuracy as low as 5 ppm, with a moderate acquisition speed (maximum acquisition rate 20 s^{-1}), and linearity range of approximately three orders of magnitude;

2. GC coupled to high-speed MS instruments (HS)TOF (maximum acquisition frequency 500 s^{-1}), with unit-mass resolution and linearity of approximately four orders of magnitude; high-speed TOF MS are in particular useful for detection and quantitative analysis in fast and ultra-fast GC, or comprehensive 2D-GC.

High sensitivity in full-scan mode, together with the high mass resolution power and accuracy provided, make GC–TOF MS very attractive in qualitative analysis, especially for screening purposes and identification of unknowns.[34,35]

Liquid Chromatography (LC-UV-DAD, LC-MS)

Liquid chromatography (LC) is the technique of choice for non-volatile compounds within a wide range of polarities, and it is very popular in foodomic studies involving proteins and non-volatile metabolites.
The main separation modes in HPLC are (i) normal phase (NP) and (ii) reverse phase (RP). In NP, the stationary phase is polar and the analytes are separated by elution with an ever *increasing* polarity mobile phase, while, in RP, the stationary phase is non-polar and an ever *decreasing* polarity such as water or acetonitrile is used.
A high number of –omic studies focuses on food identitation, classification, authentication and properties prediction through the phenolic/flavonoidic fraction. These fractions are deeply involved in many interesting properties of foods mainly because of their antioxidant activity and their bitter taste and astringent sensations.
In literature the phenolic and flavonoid fractions have been studied on several matrices such as green tea[34,36–40], red wine[41,42], honey[43] and, of course, coffee[44], by coupling HPLC (High Pressure Liquid Chromatography) to a UV/DAD as detector. UV detector suffers of some limitations, in particularly for analytes without UV chromophores, but it is the best compromise between sensitivity, linearity, versatility, and reliability versus costs.[45]
Many foodomic target analytes absorb UV light in the range of 200–550 nm, including substances with conjugated double bonds and/or un-shared and easy to excite electrons.
Three types of UV detectors are available: fixed wavelength, multiple wavelength, or photodiode array (DAD). The fixed-wavelength detector is the cheapest and has high intrinsic sensitivity because the lamp emits light at specific wavelengths. The multiple-wavelength, although less sensitive, is more versatile since it can operate at different wavelength during the same chromatographic run. UV-DAD provides the full UV spectrum of each peak and in combination with HPLC offers for the UV characterization of all separated analytes with diagnostic chromophores, (e.g. phenolics, polyketides, alkaloids, and terpenoids[46]). DAD-UV detection records the absorbance at all wavelengths of an analyte simultaneously. Limits of detection (LODs) can reach 10^{-8}g/mL, with a linear dynamic range of about three orders of magnitude.

HPLC-MS is a key technique for the online identification of food components. Besides detection, mass spectrometry gives the possibility of generating either nominal mass molecular ions, or accurate mass measurements for the determination of empirical formulas[45].
An LC–MS system includes the autosampler, the HPLC system, the ionization source (which interfaces the LC to the MS) and the mass analyzer.
It should be noted that coupling an HPLC with MS is more complex than coupling it with GC because HPLC operates with a liquid mobile phase at atmospheric pressure while MS spectrometer works in gas phase under high vacuum. This condition imposes some restrictions to the mobile phase composition and flow rate.
A number of interfaces/ionization sources have been developed. The most popular are (i) electrospray ionization (ESI) and (ii) atmospheric pressure chemical ionization (APCI); both interfaces are nowadays available in LC–MS systems. For both ESI and APCI, the ionization occurs at atmospheric pressure and is obtained by a combination of high voltage and heat.[47]
In ESI, the high voltage (3–5 kV) produces nebulization of the HPLC effluent resulting in charged droplets that are focused toward the mass analyzer. These droplets get ever smaller because of solvent evaporation as they approach the entrance to the mass analyzer forming individual ions (the so called ‘ion-evaporation’) These ions are then separated in the MS analyzer.
In APCI, heat vaporizes the column effluent and then a corona discharge is used to ionize solvent molecules, which then produce the analyte ions via chemical ionization mechanisms.

More recently, a third ionization mode, atmospheric pressure photoionization (APPI), has become available. In APPI, heat is used to vaporize the column effluent (similarly to APCI), but ionization is produced with an ultraviolet (UV) lamp that produces 10 eV photons. APPI methods in general provide a soft ionization and mainly molecular ion species in the form of either protonated molecules $[M+H]^+$ (positive-ion mode, PI) or deprotonated molecules $[M-H]^-$ (negative-ion mode, NI). Different adducts (e. g., $[M+Na]^+$ (PI) or $[M+HCOO]^-$ (NI)) are also produced, depending on solutes and modifiers.

Recently some promising results has been achieved by Termopoli et al. that developed an interface combining a conventional HPLC system to a conventional Electron Ionization (EI) mass spectrometer [48]. This new interface called liquid-EI (LEI) aims to fill the lack of structural information gap with ESI ion source which are of high interest for –omic approaches in food analyses [48].

Several analyzers are available for interfacing with HPLC:

- Single quadrupole mass analyzer; this system will provide a mass spectrum (or better molecular or quasi-molecular ion(s)) for each chromatographic peak eluting from the LC column.
- high resolution time-of- flight (HR-TOF) mass analyzer, which also provides exact masses at four to five decimal figures of each eluting component.
- Triple-quadrupole (QQQ) MS-MS systems are is the most commonly used in particular for bioanalysis and metabolite identification assays or specific quantification in complex matrices through Multiple Reaction Monitoring (MRM).
- HPLC-ion-trap mass spectrometry which has the unique capability of producing MS^n data that are important in structure elucidation studies

In addition, a growing number of additional analyzers, including hybrid systems are now available. Hybrid mass spectrometers combine two of the basic MS analyzers to make a specialty system; an example of a hybrid mass spectrometer is the 'Q-TOF' MS–MS system, which combines a quadrupole analyzer with a TOF analyzer.

High resolution mass spectrometry (HRMS) improved the field of application of LC-MS technique affording to discriminate compounds on their exact masses and gives new tools to investigate complex samples[49].
Among HRMS analyzers coupled to HPLC and UHPLC systems [(magnetic sector, time-of-flight (TOF), Orbitrap, and Fourier transform ion cyclotron (FT-ICR)], TOF and Orbitrap are the most commonly used in -omic applications.
Compared to TOF analyzers the Orbitrap has an advantage given by the relative stability of its mass calibration.
The high accuracy on mass measurements achieved with this kind of instruments enhances the possibility to unambiguously determine the elemental composition of known and new constituents with a high level of accuracy. A recent study on polyphenols in food[50] and the foodomic study of mycotoxins[51] are examples in witch this informative power has been fully exploited.

Multidimensional Chromatographic techniques

The important increase in hyphenated separations supports the idea that more information is needed to decipher the high complexity of food samples and their real effects on human health.
A measure of sample complexity is given by the parameter "sample dimensionality s" introduced by Giddings at the end of the 80[s] [52–54] which corresponds to the number of independent variables that have to be defined to decrypt the sample composition. In his studies Giddings demonstrates that as much system dimensionality (n) and the sample dimensionality (s) are close one another as much the separation pattern is ordered and its complexity resolved.
Multidimensional chromatography is a precious tool in the analysis of complex samples (as foods) because, improving system dimensionality (n), it provides a peak capacity enhancement by far higher than any improvement or optimization on mono-dimensional separations.
A common compositional characteristic of complex food samples is the highly variable abundance [from traces (ng/g) to several percent (g/100 g)] of their components, which mainly consists of groups of chemically correlated compounds such as hydrocarbons, alcohols, carbonyl derivatives, acids, and esters. In addition, when technological processes are applied to transform a raw material into a food end-product, sample compositional complexity increases.
A clear example is the roasting process, applied to several raw materials such as green coffee, cocoa beans, and hazelnuts, that must be carefully monitored and controlled in order to obtain products corresponding to the desired standards of quality [7].
These compounds sometimes show similar chromatographic retention behavior and MS fragmentation patterns that make their one-dimensional characterization and quantitation difficult. Multidimensional techniques can overcome these limitations, thanks to their increased separation power and to the possibility to adopt different separation mechanisms in the two chromatographic dimensions.
Multidimensional chromatography affords combination of two or more independent or nearly independent separation steps (first ^{1}D and second ^{2}D dimensions), increasing significantly the separation power (or more correctly the peak capacity) of the corresponding one-dimensional techniques and, therefore, the physical separation of compounds in complex samples.[1]
Multidimensional techniques can be splitted basically in heart-cutting multidimensional approaches (e.g., LC–LC, GC–GC, LC–GC, LC–CE) and comprehensive two-dimensional techniques (LC×LC, GC×GC).
Despite they both exploit the concept of orthogonality to separate analytes in complex mixture, in the heart-cut multidimensional chromatography MDGC (heart-cut 2DGC), only a limited number of ^{1}D regions (defined heart-cuts), where coelutions occur, are transferred to the second dimension (column - ^{2}D) to improve the separation[55].
Sampled fractions should have low duration to prevent the overlapping of transferred peaks on 2D column; for this reason, the conventional Heart Cut Multidimensional chromatographic (H/C MDC) experiments can be considered a limited multidimensional separation method because it can only be applied to selected portion(s) of the chromatogram. The main advantage is the possibility of an *ad-hoc* tuning of the ^{2}D column selectivity depending on each specific targeted analytical problem.

On the other hand, in a comprehensive platform (GC×GC or LC×LC), each component eluting from the first column is on-line and automatically trapped, refocused, and re-injected into a second column through a modulator, a thermal or valve-based focusing device. The time between each trapping, refocusing and re-injection events is called modulation time; it is fixed (usually between 4 and 8 s) and correspond to the the time available for the second-dimension analysis [56].

The main technical issues of multidimensional chromatography is the connection of the two dimensions; they are related, for instance, to the relatively costly operation conditions in GC×GC or to the loss in sensitivity in LC×LC.
New solutions are now available to facilitate dimension combination as well as to increase the orthogonality of the systems and, thereby, their separation power and possibility of applications to food analysis.
The development of sample preparation methods and techniques online combined with multidimensional systems and powerful MS detectors would ever more extend their use in food analysis.

The great development of two-dimensional chromatography has found many applications in the flavour and fragrance fields by providing accurate aroma fingerprints of complex samples and a better food aroma blueprints.
Multidimensional techniques may be a valuable tool for the assessment of food quality and authenticity, the control of technological processes, the determination of nutritional value, and the detection of molecules with a possible beneficial effect on human health. [22,57,58,59,60,61,62]
This full samples characterization can be exploited for many purposes like food design and off-flavours unveiling.
Multidimensional techniques have also been used to study the impact of particular food metabolites on health, to detect possible markers of diet-induced metabolic derangements or metabolic diseases on biological fluids [63,64,65,66–68,69].

Both scientific and technological world make important efforts to make the multidimensional techniques simpler, cheaper and faster[70] since these multidimensional techniques require dedicated laboratories, equipment, and highly trained personnel to be properly managed.
Despite the technological progress makes the huge amount of information provided by multidimensional techniques easier to handle this full characterization (and moreover all the related complexity) is not always needed in particular for those applications that need to be implemented on a routine-based scenario.

Direct mass spectroscopy techniques and Electronic noses

In recent years, the development of rapid, direct mass spectrometric techniques has achieved an increasing interest in flavour analysis due their potential to study flavour release , while, the conventional electronic noses (e-noses) simulate the human olfactory system by an array of chemical sensors able to screen (without providing a chemical information) a rather wide range of odorants.[22]

These techniques have in common the absence of a previous separative step that makes the statistical data elaboration (usually performed by chemometric tools) ever more crucial to decode the resulting signal.

The mains advantages of non-separative techniques are the short analysis time and the possibility of an on-line monitoring of the samples headspace. Several requirements are needed from an instrument for a reliable real-time analysis in a flavour release study, among others high sensitivity for detection of aroma markers and a fast analytical response time without the need for sample pre-treatment. Among others, direct MS techniques have become the most promising because of their signal stability, higher flexibility and information potential.

Several approaches have been used in this respect, including both the sampling techniques such as SHS and HS-SPME and mass spectrometric techniques such as proton transfer reaction mass spectrometry (PTR-MS)[71], atmospheric pressure chemical ionization mass spectrometry (APCI-MS) and selected ion flow tube mass spectrometry (SIFT-MS)[72]. Another direct MS technique with a great potential in real-time flavour analysis is Direct Analysis in Real Time Mass Spectrometry (DART-MS), where molecules from the sample (solid, liquid, or gas) are ionized by a beam of neutral metastable species generated by a glow discharge plasma. The introduction of TOF analyzer positively contributed in terms of speed and specific structural information and opened the non separative mass-based techniques field to many different applications.

Electronic noses (e-noses) are different from most other instruments used in flavour analyses since they are mainly designed to recognize gas mixtures as a whole without identifying individual chemical species within the mixture. This technique codifies any odor stimulus as a distinctive electronic pattern or fingerprint, which is further classified by comparisons with reference electronic patterns in a database. In spite of some drawbacks (e.g. sensor poisoning, sensitivity to moisture, poor linearity, etc.), over the recent years, e-noses found many applications in food industries mainly in objective quality assessments, off-flavour detection etc. [73,74].

Non-separative MS methods, better known as mass spectrometry-based electronic nose (MS-EN), were introduced by Marsili[75] to study off-flavours in milk; they have since successfully been applied to characterizing several matrices in the food field.

They provide a representative, diagnostic, and generalized mass spectrometric fingerprint of the volatile fraction of a sample, directly analyzed without prior chromatographic separation, in which each m/z ratio acts as a “sensor”, whose intensity derives from the contribution of each compound producing that fragment. These methods, in combination with appropriate chemometric elaboration, can be used to quickly characterize and discriminate samples within a set and to correlate them with a technological process, specific geographic origin or sensory characteristics. MS-EN can also be used to monitor target compounds in a group of samples, provided that specific and diagnostic ions are obtained with a compatible ion generation mode (EI, CI, APCI, PTR, etc.).

Examples of interesting applications of MS-EN on coffee aroma studies are those published in 2008 by our group on monitoring of the coffee roasting degree [76] and by Lindinger et al., who predicted the coffee sensory profiles by PTR-MS[71].

Chemometrics and Data Elaboration a crucial step of foodomic studies

Applications of foodomics include studies of foods for control, safety, functionality, authenticity, traceability, freshness, contaminants, toxicity, bioactivity and effects on human health.
Foodomics studies are thus often designed to investigate well-defined scientific questions related to food and its consumption.[77]
Modern platforms provide with an ever-increasing number of complementary signals that can be measured within a single run. Datasets become not only gradually larger but also structurally more complex, and the information from hundreds or even thousands of metabolites together with the natural biological variation of individuals make exploration of the data not-easy.
Challenges related to the use of this wealth of data include: i) the extraction of relevant elements within massive amounts of signals possibly spread across different tables, ii) reduction of dimensionality, and iii) summarization of the information in an understandable and easy way to display for interpretation purposes. The complexity of a chemical system necessarily affects the intrinsic complexity of the system itself. Foodomics thus cannot be developed without taking into account the current methodologies for exploring and exploiting patterns and relations in the mega/multivariate set of data that are generated.

For these reasons, the last but not least part of a TAS mainly concerns the data elaboration tools which, are becoming at least important as the sampling/analytical techniques in the achievement of the final results of a foodomic (or more in general omic) study.
Chemometric or multivariate data analysis are mathematical-statistical tools that try to separate "useful" information from all those contained in the data set, i.e. experimental noise, redundant information due to the correlation of some variables, information of good quality, but not directly interesting for the studied problem. Therefore, to face problems of "high complexity", it is necessary to extract the relevant information from the experimental data.
Multivariate data structures are included in tables of numbers (the data matrix) consisting of a number of observations, each one represented by variables that describe the observations.
A common data table is in general a matrix, whose *nx* rows represent the objects (samples, experiments, etc.) and the *ny* columns indicate the variables describing each object (figure 1.3.10).

The variables can be distinguished into two logical groups: X predictors (the independent variables) and the Y block of the responses (the dependent variables). Depending on the problem dealt with, a variable can belong to the block X or to the block Y.
Objects can also be associated to a vector, based on a predefined criterion, which contains the information of each object belonging to one of the predefined G classes (categories, groups, etc.). In other words, each object corresponds to a number, which identifies the class to which it belongs.
The variables are the quantities used to study a given phenomenon and to describe the observations as a whole and can be theoretical and/or experimental.
The objects are the samples analyzed to understand the investigated phenomenon, to build the models, and to confirm the hypotheses formulated. A sample in foodomics is represented by several experimental measures: the objects are here defined as multivariate. The set of measurements performed on a sample is featured by the variables selected to describe the object; the set of values defining it is the datum. In many cases, the available samples are not homogeneous since they come from different populations or belong to different classes or categories. The practical limit is the presence of a significant number of objects able to describe each class adequately.

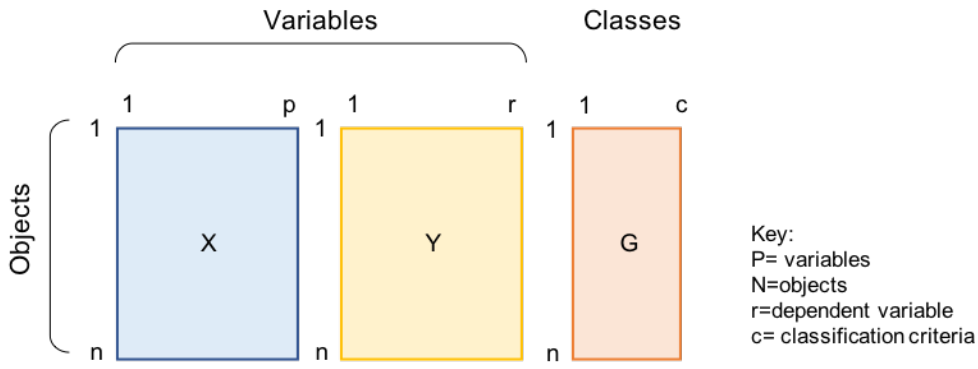


Figure 1.3.10 scheme of a common data matrix

Chemometric tools can be classified into two main groups (i) unsupervised methods, they are exploratory methods just showing the data as they are used to visualize the natural distribution of a sample in a three-dimensional space, and (ii) supervised methods, they look for determined features within data, explicitly oriented to address particular issues. In particular, when a model is developed with the purpose of predicting a qualitative or quantitative property of interest, its reliability in prediction should be assessed prior to using the model in practice. Prediction ability values should be presented together with their confidence interval, which depends on the number of samples used for the validation. The most common validation strategies divide the available samples into two subsets: a training (or calibration) set used to calculate the model and an evaluation set used to assess its reliability. The estimation of the predictive ability on new samples, not used for building the models, is a fundamental step in any modeling process and several procedures have been deployed to this purpose. No information from the test set can be used in building the model and in the pre-processing stages, otherwise the prediction ability may be overestimated. [78,79]
In several modeling techniques, some parameters are optimized looking for a setting that provides the maximum predictive ability for the model for a given sample subset. In the next paragraph, some of these parameters are discussed in function of the chemometrics tools and purposes (i.e. exploratory data analysis or prediction).

The purpose of unsupervised learning is to summarize, explore and discover encrypted data information without any a priori knowledge. Unsupervised learning is usually the first step in data analysis and can help to visualize the data or to verify any intrinsic relationship in the data matrix.
Among the unsupervised learning methods, two common methods used in metabolomics data analysis are here described and discussed:

- Principal Component Analysis (PCA);
- Multiple Factor Analysis (MFA)

Unsupervised data elaboration

Principal Component Analysis (PCA)

PCA is one of the most powerful tools to perform datasets reduction; it allows to find only a few combinations of the original variables (those measured in the experiments) that best explain the total variation in the original dataset.
PCA can be considered as the starting point of multivariate data analyses; its main objective is to replace all correlated variables by a much smaller number of uncorrelated variables, often referred as Principal Components (PCs), that still retain most of the information in the original dataset.
In practice, PCA builds hyperplanes in the original feature space that are linear combinations of the original variables (those measured in the experiments) and describes the data according to these new variables. The inspection of PCA scores highlights the relationships among the distribution of samples that may reveal groups while loading plots encloses the variables describing the groupings.
The data can be visualized with a two-dimensional or three-dimensional plot (called scores and loadings plots) in which the large proportion of variation in the data is explained.
The first variable (Principal Component 1, PC1) must have the largest possible variance to cover the largest amount of variability inside of the data set. The second principal component (PC2) must be orthogonal to the first one in order to explain the larger amount of the remaining variance. The other components are computed likewise.
There is no rule for how many PCs to keep; the decision is usually made by checking the ‘‘variance explained’’ measure mentioned above or using a screen plot.
Each observation is described on the new variables by values that are called factors scores; these values can be interpreted as the projections of the observation onto each principal component.
PCA elaboration provides many outputs but, for data exploration purposes the most important information are reported as follow:

- **Total Explained Variance**: since PCA provides a reduction of data dimension, it is important to verify that the amount of variance explained by the PCs considered is sufficient to explain and summarize the phenomenon under study.
- **Pearson’s Correlation Matrix**: this matrix contains the Pearson’s correlation coefficient calculated between one original variable and each of the other variables. The Pearson’s correlation coefficient ranges between 0 and 1 and indicates the degree of correlation between two variables; a high correlation between two variables mean that they provide the same information.
- **Scores Plot**: it is a 2 or 3D plot where each observation is plotted *versus* the two or three principal components chosen. This diagram is particularly useful to highlight if the pool of observations is grouped or clusterized by PCs.
- **Loadings Plot**: this diagram plots the original variables *versus* the PCs. The projection of each variable on the PCs indicates the weight of the original variables in the definition of score samples. While the scores plot is useful to observe how the samples are scattered on the PCs space, the loadings plot affords to point out the original variables mostly responsible for this scattering.

Figure 1.3.11 represents PCA workflow: objects (rounds and triangles) are described by variables (X_1, X_2 and X_3). The elaboration provides two plots; the Scores and the Loadings plots where the objects and the variables are plotted according to the PCs selected (t_1 and t_2). In this simple scheme, the interpretation is quite simple; circles are almost well separated from triangles according to the first PC (here called t_1). The Loadings plot show that circles are characterized by the variable X_3 while the triangles by variables X_1 and X_2.

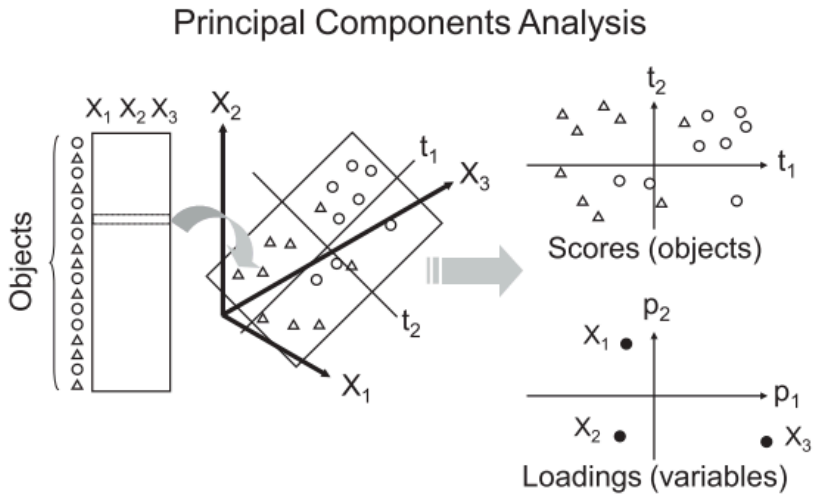


Figure 1.3.11 PCA elaboration scheme

Multiple Factor Analysis (MFA)

This unsupervised analytical tool, introduced by Escofier and Pages at the beginning of the 90[th][80], affords to analyse simultaneously several tables of variables, and to obtain results, in particular charts, to study the relationship between the observations, the variables and the tables.
It takes into account the contribution of all active groups of variables to define the distance between individuals. The number of variables in each group may differ and the nature of the variables (qualitative or quantitative) can vary from one group to the other but the variables should be of the same nature in a given group. The goal of MFA is to integrate different groups of variables describing the same observations[81]. In MFA analyses the observations are described by several “blocks" or sets of variables. MFA seeks for the common structures present in all or some of these sets. MFA is performed in two steps. First, a principal component analysis (PCA) is performed on each data set that is then “normalized” by dividing all its elements by the square root of the first eigenvalue obtained from of its PCA. Second, the normalized data sets are merged to form a unique matrix and a global PCA is performed on this matrix[82]. The individual data sets are then projected onto the global analysis to analyse communalities and discrepancies. The weighting of the tables makes it possible to prevent that the tables including more variables not weighting too much in the analysis. MFA is used in very different domains such as sensory evaluation, economy, ecology, and chemistry.

The most interesting information obtained by the MFA are listed below:

- Within the global elaboration, the contribution of each group of variables on the PCs is reported in form of percentage;

- **LG Coefficients:** the LG coefficients measures how the tables are related two by two. The more variables of a first table are related to the variables of the second table, the higher the LG coefficient.
- **RV Coefficients**: these values are more intuitive than LG coefficients; they are related to LG coeff. but, for an easier interpretation they range between 0 and 1. This value can be assumed as a generalization of the Pearson's correlation coefficient.

The potentials of this data elaboration technique have been exploited during section 2.4 of the experimental part, where some Aroma and Taste features have been compared in their consistency in describing the pool of samples (observations) under study.

Supervised data elaboration tools

Supervised learning methods are widely used to discover biomarkers, classification, and prediction.
Supervised learning deals with problems or datasets with response variables. These variables can be either discrete or continuous. When the variables are discrete, e.g., control group versus diseased group, the problems are called *classification problems*. When the variables are continuous, e.g., metabolite concentration or gene expression level, the problems are called *regression problems*. The purpose of supervised learning is to determine the association between the response variable and the predictors (often referred to as covariates) and to make accurate predictions. This operation is called *supervised learning* because one or more response variables are used to guide the training of the models. Usually a training and an evaluation step are included while building and test the fit of the model, and then the testing dataset is used to evaluate the predictive power[77].

Partial Least Square (PLS) Regression

PLS regression is a recent technique that generalizes and combines features from principal component analysis and multiple regression. In particular, it is useful when we need to predict a set of dependent variables from a (very) large set of independent variables (i.e., predictors).
Partial least squares modeling is a multivariate projection method for modeling a relationship between dependent variables (Y) and independent variables (X). The principle of PLS is to find the components (similar to those calculated in the PCA) in the input matrix (X) that describe as much as possible the relevant variations in the input variables, and, at the same time, have maximal correlation with the target value in Y, giving less weight to the irrelevant or noisy variations
PLS therefore simultaneously models both X and Y to find the latent variables in X that will predict the latent variables in Y. PLS maximizes the covariance between matrices X and Y. An important feature of PLS is that it takes into account errors in both matrices, X and Y, and assumes that they are equally distributed. Moreover, PLS is suitable for data sets with fewer objects than variables and a high degree of inter-correlation between the independent variables [78].
When the Y matrix is formed by qualitative variables, the PLS is used for discrimination purposes and takes the name of Partial Least Square Discriminant Analysis (PLS-DA).[83]
The partial least squares discriminant analysis aims to find the variables and directions in the multivariate space which discriminate the established classes in the calibration set. The optimal number of latent variables can be estimated by an internal cross-validation or an evaluation test sets. In order to obtain a better model, it is important to work with a well-designed training set; in other words, it has to be representative of the phenomenon under study both in terms of variability and numerousness. The test set is a set of objects in which the dependent variable (Y)

has to be predicted by the PLS model. This dataset has to be totally external from the built model, i.e. it never has to be used in the model training steps.

Most common parameters used to evaluate PLS regression model performance are reported as follow:

- **Q^2 index**; this is a model quality index, this value measures the global contribution of the components to the predictive quality of the model. The search of the maximum value of Q^2 is equivalent to finding the most stable model.
- **Determination Coefficient of the model (R^2)**; this coefficient, whose value ranges between 0 and 1, indicates the proportion of variability of the dependent variable explained by the model. The nearest to 1 R^2 is the better is the model.
  The main issues of R^2 is that it does not take into account the number of variables used to fit the model. Adjusted R^2 can be used instead of R^2 to overcome this limit. The number of variables used for the model development is important since the number of unnecessary variables penalizes the model; Adjusted R^2, unlike R^2, is sensitive to these penalties. Adjusted R^2 can be calculated with the following formula:

$$_{\text{Adj}}R^2 = 1 - (1 - R^2) \times \frac{n-1}{n-p}$$

  where R^2 is the determination coefficient of the model, n and p are the number of observations and variables used to fit the model.
- **Root Mean Squared Error (RMSE)** is defined as the square root of the Mean Squared Error that measures the average of the squares of the errors or deviations, that is the difference between the predicted values and what is measured. When the error is calculated on cross validation data takes the name of root mean squared error in Cross Validation (RMSECV) while when it is calculated on prediction data root mean squared error in prediction (RMSEP).
-

$$RMSE = \sqrt{\frac{\sum_{I=1}^{n} w_i \, (y_i - \hat{y}_i)^2}{W - p^*}}$$

Non-linear and non-parametric regressions

Although the relationship between dependent and independent variables is often assumed as linear, this assumption does not always provide the best prediction models. Non-linear and non-parametric regression algorithms are therefore successfully used, in particular for biological phenomena as those monitored in foodomics.
Examples of non-linear regression have been applied and developed by Abdel-Rahman et al in 2009 [84] and by Yoo et al. in 2004 [85]. In the first paper the Takagi–Sugeno–Kang (TSK) fuzzy inference system has been combined to PLS to overcome its limits in terms of flexibility when non-linear relationships are involved. The Quadratic Fuzzy PLS (QFPLS) have been developed from the combination of the two approaches. A series of experiments carried out in the authors' laboratory on several on data of different nature and sizes have shown that QFPLS clearly outperformed four other well-known methods (Linear Partial Least Square (LPLS), Quadratic Partial Least Square

(QPLS), Fuzzy Partial Least Square (FPLS), and Neural Net Partial Least Square (NNPLS) described in the literature.
Kernel regression and LOWESS regression are part of *non-parametric regressions*.
These approaches can be used when the hypotheses about the classical regression methods cannot be verified or when the interest is mainly focused only on predictive quality of the model and not on its structure (i.e. its equation). They are based on polynomial regression using and use three subsets of samples as already described before in the building up and testing the model.

The non-parametric regression algorithm most frequently used is known as "LOWESS" (LOcally WEighted regression and Smoothing Scatter plots) curve. The acronym represents the idea of a locally weighted regression curve whose values, at a specific location along the x-axis, is determined by the points in that vicinity. As a consequence, the method does not imply any assumptions about the form of the relationship and affords to discover the trend using the data itself.
This regression is in particular indicated when the interest is to improve the predictive quality and not to investigate the structure of the model. The lack of information about the structure of the models and their equation are the main limits to its application[1].
This last algorithm has been used to optimize the chemical fingerprints of the single notes in the last session of the experimental part (2.5).

[1] (https://help.xlstat.com/customer/en/portal/articles/2062253-nonparametric-regression-kernel-lowess-tutorial?b_id=9283)

References

1. García-Cañas, V., Simó, C., Herrero, M., Ibáñez, E. & Cifuentes, A. Present and Future Challenges in Food Analysis: Foodomics. *Anal. Chem.* **84,** 10150–10159 (2012).
2. Capozzi, F. & Bordoni, A. Foodomics: A new comprehensive approach to food and nutrition. *Genes Nutr.* **8,** 1–4 (2013).
3. Gallo, M. & Ferranti, P. The evolution of analytical chemistry methods in foodomics. *J. Chromatogr. A* **1428,** 3–15 (2016).
4. Fallis, A. . & Lynch, S. Summary for Policymakers. in *Climate Change 2013 - The Physical Science Basis* (ed. Intergovernmental Panel on Climate Change) **53,** 1–30 (Cambridge University Press, 2013).
5. West, J., Becker, M., Tombrink, S. & Manz, A. Micro total analysis systems: Latest achievements. *Anal. Chem.* **80,** 4403–4419 (2008).
6. Manz, A., Graber, N. & Widmer, H. M. Miniaturized total chemical analysis systems: A novel concept for chemical sensing. *Sensors Actuators B Chem.* **1,** 244–248 (1990).
7. Pico, Y. *Chemical Analysis of Food Techniques ans Application*. *Elsevier* (2012).
8. Buettner, A. *Handbook of Odor*. (Springer, 2017).
9. Bicchi, C., Cordero, C., Liberto, E., Sgorbini, B. & Rubiolo, P. Headspace Sampling in Flavor and Fragrance Field. in *Comprehensive Sampling and Sample Preparation* **4,** 1–25 (Elsevier, 2012).
10. Collin, S., Nizet, S., Muls, S., Iraqi, R. & Bouseta, A. A. Characterization of odor-active compounds in extracts obtained by simultaneous extraction/distillation from Moroccan black olives. *J. Agric. Food Chem.* **56,** 3273–3278 (2008).
11. Engel, W., Bahr, W. & Schieberle, P. Solvent assisted flavour evaporation - a new and versatile technique for the careful and direct isolation of aroma compounds from complex food matrices. *Eur. Food Res. Technol.* **209,** 237–241 (1999).
12. Souza-Silva, É. A., Gionfriddo, E. & Pawliszyn, J. A critical review of the state of the art of solid-phase microextraction of complex matrices II. Food analysis. *TrAC Trends Anal. Chem.* **71,** 236–248 (2015).
13. Cordero, C. *et al.* High concentration capacity sample preparation techniques to improve the informative potential of two-dimensional comprehensive gas chromatography–mass spectrometry: Application to sensomics. *J. Chromatogr. A* **1318,** 1–11 (2013).
14. Tienpont, B., David, F., Bicchi, C. & Sandra, P. High capacity headspace sorptive extraction. *J. Microcolumn Sep.* **12,** 577–584 (2000).
15. Bicchi, C. *et al.* Dual-phase twisters: A new approach to headspace sorptive extraction and stir bar sorptive extraction. *J. Chromatogr. A* **1094,** 9–16 (2005).
16. Sgorbini, B. *et al.* Solvent-enhanced headspace sorptive extraction in the analysis of the volatile fraction of matrices of vegetable origin. *J. Sep. Sci.* **33,** 2191–2199 (2010).
17. Snow, N. H. & Bullock, G. P. Novel techniques for enhancing sensitivity in static headspace extraction-gas chromatography. *J. Chromatogr. A* **1217,** 2726–2735 (2010).
18. Kolb, B. & Ettre, L. S. *Static Headspace-gas chromatography: Theory and Practice*. *Zhurnal Analiticheskoj Khimii* **57,** (2002).
19. Bicchi, C., Cordero, C. & Rubiolo, P. A survey on high-concentration-capability headspace sampling techniques in the analysis of flavors and fragrances. *J. Chromatogr. Sci.* **42,** 402–9 (2004).
20. Bicchi, C. *et al.* Quantitative analysis of volatiles from solid matrices of vegetable origin by high concentration capacity headspace techniques: Determination of furan in roasted coffee. *J. Chromatogr. A* **1218,** 753–762 (2011).
21. Bicchi, C., Cordero, C., Iori, C., Rubiolo, P. & Sandra, P. Headspace Sorptive Extraction (HSSE)

in the Headspace Analysis of Aromatic and Medicinal Plants. *J. High Resolut. Chromatogr.* **23,** 539–546 (2000).

22. Regueiro, J., Negreira, N. & Simal-Gándara, J. Challenges in Relating Concentrations of Aromas and Tastes with Flavour Features of Foods. *Crit. Rev. Food Sci. Nutr.* **8398,** 0 (2015).
23. Mcgorrin, R. J. One Hundred Years of Progress in Food Analysis. *J. Agric. Food Chem.* **57,** 8076–8088 (2009).
24. Cuadros-Rodríguez, L., Ruiz-Samblás, C., Valverde-Som, L., Pérez-Castaño, E. & González-Casado, A. Chromatographic fingerprinting: An innovative approach for food 'identitation' and food authentication – A tutorial. *Anal. Chim. Acta* **909,** 9–23 (2016).
25. Jiao, J. *et al.* Study of Chromatographic Fingerprint of the Flavor in Beer by HS-SPME-GC. *Anal. Lett.* **44,** 648–655 (2011).
26. Jacobson, D., Monforte, A. R. & Ferreira, A. C. S. Untangling the Chemistry of Port Wine Aging with the Use of GC-FID, Multivariate Statistics, and Network Reconstruction. *J. Agric. Food Chem.* **61,** 2513–2521 (2013).
27. Ochi, H., Bamba, T., Naito, H., Iwatsuki, K. & Fukusaki, E. Metabolic fingerprinting of hard and semi-hard natural cheeses using gas chromatography with flame ionization detector for practical sensory prediction modeling. *J. Biosci. Bioeng.* **114,** 506–511 (2012).
28. Jumtee, K., Bamba, T. & Fukusaki, E. Fast GC-FID based metabolic fingerprinting of Japanese green tea leaf for its quality ranking prediction. *J. Sep. Sci.* **32,** 2296–2304 (2009).
29. Durante, C., Cocchi, M., Grandi, M., Marchetti, A. & Bro, R. Application of N-PLS to gas chromatographic and sensory data of traditional balsamic vinegars of modena. **83,** 54–65 (2006).
30. Ribeiro, J. S., Augusto, F., Salva, T. J. G. & Ferreira, M. M. C. Prediction models for Arabica coffee beverage quality based on aroma analyses and chemometrics. *Talanta* **101,** 253–260 (2012).
31. Pollien, P. *et al.* Hyphenated Headspace-Gas Chromatography-Sniffing Technique: Screening of Impact Odorants and Quantitative Aromagram Comparisons. *J. Agric. Food Chem.* **45,** 2630–2637 (1997).
32. Purcaro, G. *et al.* Evaluation of a Rapid-Scanning Quadrupole Mass Spectrometer in an Apolar × Ionic-Liquid Comprehensive Two-Dimensional Gas Chromatography System. *Anal. Chem.* **82,** 8583–8590 (2010).
33. Hernández, F. *et al.* Current use of high-resolution mass spectrometry in the environmental sciences. *Anal. Bioanal. Chem.* **403,** 1251–1264 (2012).
34. Daszykowski, M., Vander Heyden, Y. & Walczak, B. Robust partial least squares model for prediction of green tea antioxidant capacity from chromatograms. *J. Chromatogr. A* **1176,** 12–18 (2007).
35. Binkley, J. & Libarondi, M. Comparing the Capabilities of Time-of-Flight and Quadrupole Mass Spectrometers. *LCGC* **8,** 28–33 (2010).
36. van Nederkassel, A. M., Daszykowski, M., Massart, D. L. & Vander Heyden, Y. Prediction of total green tea antioxidant capacity from chromatograms by multivariate modeling. *J. Chromatogr. A* **1096,** 177–186 (2005).
37. Dumarey, M., van Nederkassel, A. M., Deconinck, E. & Vander Heyden, Y. Exploration of linear multivariate calibration techniques to predict the total antioxidant capacity of green tea from chromatographic fingerprints. *J. Chromatogr. A* **1192,** 81–88 (2008).
38. Dumarey, M., Smets, I. & Vander Heyden, Y. Prediction and interpretation of the antioxidant capacity of green tea from dissimilar chromatographic fingerprints. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **878,** 2733–2740 (2010).
39. Zhao, Y. *et al.* Tentative identification, quantitation, and principal component analysis of

green pu-erh, green, and white teas using UPLC/DAD/MS. *Food Chem.* **126,** 1269–1277 (2011).
40. Alaerts, G. *et al.* Similarity analyses of chromatographic fingerprints as tools for identification and quality control of green tea. *J. Chromatogr. B* **910,** 61–70 (2012).
41. Beltrán, N. H. *et al.* Feature extraction and classification of Chilean wines. *J. Food Eng.* **75,** 1–10 (2006).
42. Serrano-Lourido, D., Saurina, J., Hernández-Cassou, S. & Checa, A. Classification and characterisation of Spanish red wines according to their appellation of origin based on chromatographic profiles and chemometric data analysis. *Food Chem.* **135,** 1425–1431 (2012).
43. Zhou, J. *et al.* Floral classification of honey using liquid chromatography–diode array detection–tandem mass spectrometry and chemometric analysis. *Food Chem.* **145,** 941–949 (2014).
44. Minamisawa, M., Yoshida, S. & Takai, N. Determination of biologically active substances in roasted coffees using a diode-array HPLC system. *Anal. Sci.* **20,** 325–8 (2004).
45. Wolfender, J.-L. HPLC in Natural Product Analysis: The Detection Issue. *Planta Med.* **75,** 719–734 (2009).
46. Robbins, R. J. Phenolic Acids in Foods: An Overview of Analytical Methodology. *J. Agric. Food Chem.* **51,** 2866–2887 (2003).
47. Korfmacher, W. A. Foundation review: Principles and applications of LC-MS in new drug discovery. *Drug Discov. Today* **10,** 1357–1367 (2005).
48. Termopoli, V., Famiglini, G., Palma, P., Piergiovanni, M. & Cappiello, A. Atmospheric Pressure Vaporization Mechanism for Coupling a Liquid Phase with Electron Ionization Mass Spectrometry. *Anal. Chem.* **89,** 2049–2056 (2017).
49. Kaufmann, A. Combining UHPLC and high-resolution MS: A viable approach for the analysis of complex samples? *TrAC Trends Anal. Chem.* **63,** 113–128 (2014).
50. Lucci, P., Saurina, J. & Núñez, O. Trends in LC-MS and LC-HRMS analysis and characterization of polyphenols in food. *TrAC - Trends Anal. Chem.* **88,** 1–24 (2017).
51. Rychlik, M., Kanawati, B. & Schmitt-Kopplin, P. Foodomics as a promising tool to investigate the mycobolome. *TrAC Trends Anal. Chem.* **96,** 22–30 (2017).
52. Giddings, J. C. Two-dimensional separations: concept and promise. *Anal. Chem.* **56,** 1258A–1270A (1984).
53. Giddings, J. C. Concepts and comparisons in multidimensional separation. *J. High Resolut. Chromatogr.* **10,** 319–323 (1987).
54. Giddings, J. C. Sample dimensionality: A predictor of order-disorder in component peak distribution in multidimensional separation. *J. Chromatogr. A* **703,** 3–15 (1995).
55. Mondello, L., Verzera, A., Previti, P., Crispo, F. & Dugo, G. Multidimensional Capillary GC−GC for the Analysis of Complex Samples. 5. Enantiomeric Distribution of Monoterpene Hydrocarbons, Monoterpene Alcohols, and Linalyl Acetate of Bergamot (*Citrus b ergamia* Risso et Poiteau) Oils. *J. Agric. Food Chem.* **46,** 4275–4282 (1998).
56. Ryan, D. & Marriott, P. Comprehensive two-dimensional gas chromatography. *Anal. Bioanal. Chem.* **376,** 295–297 (2003).
57. Magagna, F. *et al.* Comprehensive Chemical Fingerprinting of High-Quality Cocoa at Early Stages of Processing: Effectiveness of Combined Untargeted and Targeted Approaches for Classification and Discrimination. *J. Agric. Food Chem.* **65,** 6329–6341 (2017).
58. Nicolotti, L. *et al.* Quantitative fingerprinting by headspace—Two-dimensional comprehensive gas chromatography–mass spectrometry of solid matrices: Some challenging aspects of the exhaustive assessment of food volatiles. *Anal. Chim. Acta* **798,**

115–125 (2013).
59. Magagna, F. *et al.* Black tea volatiles fingerprinting by comprehensive two-dimensional gas chromatography – Mass spectrometry combined with high concentration capacity sample preparation techniques: Toward a fully automated sensomic assessment. *Food Chem.* **225,** 276–287 (2017).
60. Purcaro, G., Cordero, C., Liberto, E., Bicchi, C. & Conte, L. Toward a definition of blueprint of virgin olive oil by comprehensive two-dimensional gas chromatography. *J. Chromatogr. A* **1334,** 101–111 (2014).
61. Mondello, L. *et al.* Comprehensive multidimensional GC for the characterization of roasted coffee beans. *J. Sep. Sci.* **27,** 442–450 (2004).
62. Cacciola, F., Donato, P., Sciarrone, D., Dugo, P. & Mondello, L. Comprehensive liquid chromatography and other liquid-based comprehensive techniques coupled to mass spectrometry in food analysis. *Anal. Chem.* **89,** 414–429 (2017).
63. Quehenberger, O. *et al.* Lipidomics reveals a remarkable diversity of lipids in human plasma,. *J. Lipid Res.* **51,** 3299–3305 (2010).
64. Bressanello, D. *et al.* Urinary metabolic fingerprinting of mice with diet-induced metabolic derangements by parallel dual secondary column-dual detection two-dimensional comprehensive gas chromatography. *J. Chromatogr. A* **1361,** 265–276 (2014).
65. Jurowski, K. *et al.* Analytical Techniques in Lipidomics: State of the Art. *Crit. Rev. Anal. Chem.* **47,** 418–437 (2017).
66. Llorach, R., Garcia-Aloy, M., Tulipani, S., Vazquez-Fresno, R. & Andres-Lacueva, C. Nutrimetabolomic Strategies To Develop New Biomarkers of Intake and Health Effects. *J. Agric. Food Chem.* **60,** 8797–8808 (2012).
67. Amann, A. *et al.* The human volatilome: volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva. *J. Breath Res.* **8,** 34001 (2014).
68. Boto-Ordóñez, M. *et al.* Microbial Metabolomic Fingerprinting in Urine after Regular Dealcoholized Red Wine Consumption in Humans. *J. Agric. Food Chem.* **61,** 9166–9175 (2013).
69. Kouremenos, K. A., Pitt, J. & Marriott, P. J. Metabolic profiling of infant urine using comprehensive two-dimensional gas chromatography: Application to the diagnosis of organic acidurias and biomarker discovery. *J. Chromatogr. A* **1217,** 104–111 (2010).
70. Reichenbach, S. E. *et al.* Alignment for Comprehensive Two-Dimensional Gas Chromatography with Dual Secondary Columns and Detectors. *Anal. Chem.* **87,** 10056–10063 (2015).
71. Lindinger, C. *et al.* When Machine Tastes Coffee: Instrumental Approach To Predict the Sensory Profile of Espresso Coffee. *Anal. Chem.* **80,** 1574–1581 (2008).
72. Biasioli, F., Yeretzian, C., Märk, T. D., Dewulf, J. & Van Langenhove, H. Direct-injection mass spectrometry adds the time dimension to (B)VOC analysis. *TrAC Trends Anal. Chem.* **30,** 1003–1017 (2011).
73. Wilson, A. Diverse Applications of Electronic-Nose Technologies in Agriculture and Forestry. *Sensors* **13,** 2295–2348 (2013).
74. Wilson, A. D. & Baietto, M. Advances in Electronic-Nose Technologies Developed for Biomedical Applications. *Sensors* **11,** 1105–1176 (2011).
75. Marsili, R. T. SPME−MS−MVA as an Electronic Nose for the Study of Off-Flavors in Milk. *J. Agric. Food Chem.* **47,** 648–654 (1999).
76. Liberto, E. *et al.* Non-separative Headspace Solid Phase Microextraction–Mass Spectrometry Profile as a Marker To Monitor Coffee Roasting Degree. *J. Agric. Food Chem.* **61,** 1652–1660 (2013).

77. Ren, S., Hinzman, A. A., Kang, E. L., Szczesniak, R. D. & Lu, L. J. Computational and statistical analysis of metabolomics data. *Metabolomics* **11,** 1492–1513 (2015).
78. Berrueta, L. A., Alonso-Salces, R. M. & Héberger, K. Supervised pattern recognition in food analysis. *J. Chromatogr. A* **1158,** 196–214 (2007).
79. Oliveri, P. & Forina, M. Data Analysis and Chemometrics. in *Chemical analysis of food: Techniques and applications* 25–57 (Elsevier, 2012).
80. Escofier, B. & Pages, J. Multiple factor analysis (AFUMULT package). *Comput. Stat. Data Anal.* **18,** 121–140 (1994).
81. Bécue-Bertaut, M. & Pagès, J. Multiple factor analysis and clustering of a mixture of quantitative, categorical and frequency data. *Comput. Stat. Data Anal.* **52,** 3255–3268 (2008).
82. Abdi, H., Williams, L. J. & Valentin, D. Multiple factor analysis: principal component analysis for multitable and multiblock data sets. *Wiley Interdiscip. Rev. Comput. Stat.* **5,** 149–179 (2013).
83. Trygg, J., Holmes, E. & Lundstedt, T. Chemometrics in Metabonomics. *J. Proteome Res.* **6,** 469–479 (2007).
84. Abdel-Rahman, A. I. & Lim, G. J. A nonlinear partial least squares algorithm using quadratic fuzzy inference system. *J. Chemom.* **23,** 530–537 (2009).
85. Yoo, C. K., Bang, Y. H., Lee, I.-B., Vanrolleghem, P. a. & Rosén, C. Application of fuzzy partial least squares (FPLS) modeling nonlinear biological processes. *Korean J. Chem. Eng.* **21,** 1087–1097 (2004).

2. Experimental Part

Coffee aroma and flavour have in depth been studied but the chemistry behind the different modulations of this complex and attractive perception is still far to be fully clarified.
The experimental part is divided into five sections in which the coffee flavour has instrumentally been studied for different purposes but always with the aim of building a bridge between the chemical characterization of coffee samples and their sensory description by a trained panel.
The final goal is to exploit this knowledge to develop a complementary tool to the time-consuming conventional sensory analysis to be used in routine controls.

Section 2.1:

The analytical approach was based on different TASs in which the sampling step has been performed through different HCC approaches and the gas chromatography-mass spectrometry (GC-qMS) has been adopted as analytical platform.
GC is the reference technique for volatiles analysis and MS provides a high level of information about sample composition; in addition, GC-MS is a relatively user-friendly and low-cost instrumentation that can usually be used in routine controls.
Since the sample preparation is the first step of the TAS and conditions the analytical information provided by the whole platform, the first approach to study this topic has been the investigation of the data obtained changing the sampling technique and their relationship with sensory data.
In order to be as close as possible to the panel experience, three different HCC samplings have been compared, each one simulating a step of the SCAA cupping protocol, (Specialty Coffee Association of America (SCAA)), officially adopted for the sensory evaluation of the incoming raw material. Chemical data handling has been performed by chemometric tools with the aim to compare the different sampling techniques to evaluate their ability to describe chemically the sensory characteristics of the samples investigated in compliance with the panel data, besides the possibility of a full automation.

Section 2.2:

The information collected in the first section have been exploited to define the chemical fingerprints of different coffee sensory notes by linking the monadic sensory profiling with their chemical composition by a Sensometric approach.
The Sensometric approach involves the use of (unsupervised and supervised) chemometric tools to explore, reduce and model both sensory and chemical data. These tools afford to consider many variables together and are at the basis of data mining in -omics studies.
Samples from all over the world have been selected and analysed, from a sensory point of view, through the standardized SCAA coffee cupping protocol to take into account the extreme variability around coffee sensory profiles. Coffee samples have here been described through seven different sensory notes.
In the so called Qualitative Discriminant Analysis, samples distribution has been explored considering only sensory data to investigate if and how the sensory profiles (defined by the scores of the panel) could discriminate the samples.
Then, the chemical data were included to select those chemicals chemometrically related to the expression of each sensory note. These compounds were at the basis of the sensory prediction models used as a validation tool of the signatures of the notes.
In agreement with these preliminary results, the sensometric approach has shown to be discriminative, informative and predictive to define the chemical signature of different aroma notes.

Section 2.3:

The main limit of the sensometric approach here adopted is the lack of a direct correspondence between the odour of the pool of compounds involved in the fingerprint of each sensory note and the sensory impact of the note itself.

The effectiveness of this strategy has therefore been compared to the molecular sensory science approach that, for two decades, has been used to identify and quantify the molecules responsible in food flavours.
In order to obtain an inter-approach validation of the fingerprints defined by Sensometrics, coffee "*woody*" and "*flowery*" sensory notes have been characterized with sensomics and sensometric approaches and the results compared. Both approaches resulted in a similar pool of compounds able to differentiate the two set of samples and thereby the expression of the "*woody*" and "*flowery"* sensory notes.
It is clear that these approaches cannot be compared without considering their substantial differences but the good consistency between the compounds pointed out with each of them candidates the sensometric approach as a valid tool to face this important challenge.

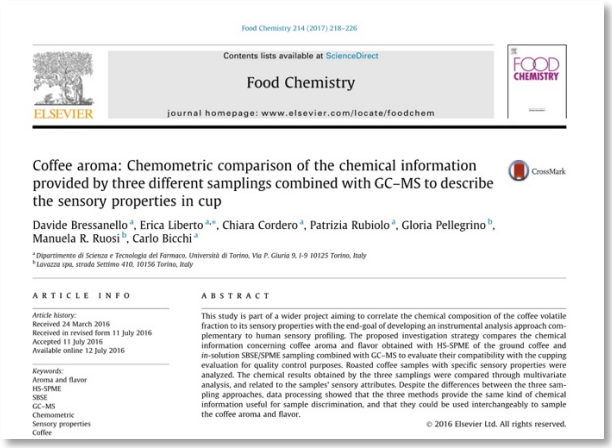
This part of the project has been carried out at the Deutsche Forschungsanstalt für Lebensmittelchemie (DFA), food chemistry institute of the TUM (Technische Universität München) under the supervision of Prof. Dr. Peter Schieberle.

Section 2.4:
The purpose of this section was to investigate in depth the complexity of coffee flavour perception to understand the limits of the sensometric approach. A Flavouromic approach has been adopted to link chemical information from volatile and non-volatile fractions to the panel flavour evaluations.
Chemical information from HS-SPME-GC-MS analysis of coffee powders was integrated with the HPLC UV/DAD fingerprinting analyses targeted on specific wavelengths of non-volatile fractions to investigate the contribution of data fusion in flavour definition and prediction.
The study is mainly focused on the *Bitter* note and then extended to the other sensory notes of interest.

Section 2.5:
The Sensometric approach discussed in section 2.2 and validated by the molecular sensory science (sec 2.3) is not only promising concerning the chemical characterization of different sensory notes but also proved to be suitable as an analytical tool complementary to the panel.
The evaluation of coffee sensory profiles, however, is highly challenging since they can assume different shades depending on several factors (origin, post-harvest-treatments, harvesting period etc.)
In this this section, the sensory notes prediction models have been optimized by a higher number of samples and with different regression algorithms able to better fit with the available data.
In addition, the models have been tested on a new test set designed to reproduce a "real-world" application condition; the sample set consists of coffee samples harvested one year later including both the same origins used to train the models together with totally new origins.

Finally, all these information have been exploited to merge all the prediction models into a single multi-note model able to predict the sensory score of all notes with one single elaboration.
The use of a single multi-note model compared to the application of six different models impose some compromises in terms of performance and stability but afford a drastic reduction of the time needed to obtain a sensory profile that is an important advantage in a routine application of the proposed strategy.

2.1 Coffee aroma: chemometric comparison of the chemical information provided by three different samplings combined with GC-MS to describe the sensory properties in cup



Food Chemistry 214 (2017) 218–226

Contents lists available at ScienceDirect

Food Chemistry

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

Coffee aroma: Chemometric comparison of the chemical information provided by three different samplings combined with GC–MS to describe the sensory properties in cup

Davide Bressanello [a], Erica Liberto [a,*], Chiara Cordero [a], Patrizia Rubiolo [a], Gloria Pellegrino [b], Manuela R. Ruosi [b], Carlo Bicchi [a]

[a] Dipartimento di Scienza e Tecnologia del Farmaco, Università di Torino, Via P. Giuria 9, 1-9 10125 Torino, Italy
[b] Lavazza spa, strada Settimo 410, 10156 Torino, Italy

ARTICLE INFO

Article history:
Received 24 March 2016
Received in revised form 11 July 2016
Accepted 11 July 2016
Available online 12 July 2016

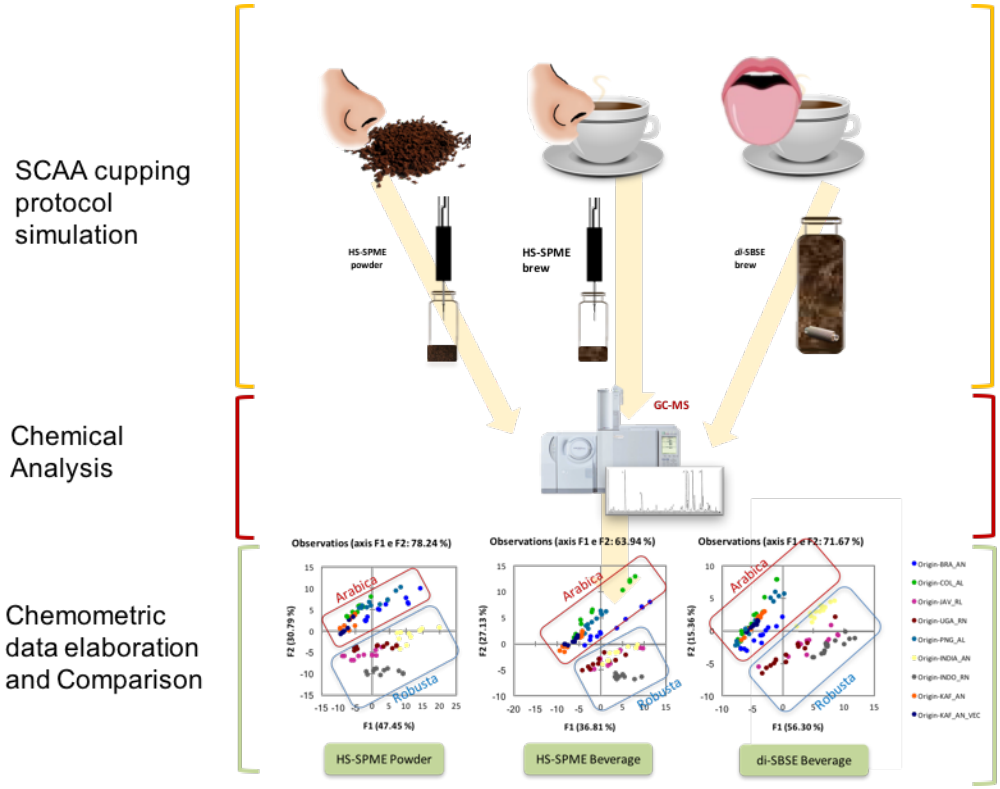
Keywords:
Aroma and flavor
HS-SPME
SBSE
GC–MS
Chemometric
Sensory properties
Coffee

ABSTRACT

This study is part of a wider project aiming to correlate the chemical composition of the coffee volatile fraction to its sensory properties with the end-goal of developing an instrumental analysis approach complementary to human sensory profiling. The proposed investigation strategy compares the chemical information concerning coffee aroma and flavor obtained with HS-SPME of the ground coffee and *in-solution* SBSE/SPME sampling combined with GC–MS to evaluate their compatibility with the cupping evaluation for quality control purposes. Roasted coffee samples with specific sensory properties were analyzed. The chemical results obtained by the three samplings were compared through multivariate analysis, and related to the samples' sensory attributes. Despite the differences between the three sampling approaches, data processing showed that the three methods provide the same kind of chemical information useful for sample discrimination, and that they could be used interchangeably to sample the coffee aroma and flavor.

This study has been published on "Food Chemistry".
DOI:https://doi.org/10.1016/j.foodchem.2016.07.088

0. GRAPHICAL ABSTRACT



1. INTRODUCTION

The quality of a cup of coffee and its distinctive sensory properties depend on the entire production chain. Some of the major factors influencing the final product are: geographical origin, climate, species, harvesting methods, technological processing (mainly roasting and grinding), storage conditions, and last but no less important, the brewing method [1,2].

Aroma and Flavour are undoubtedly important hedonic aspects of a good coffee[2], and thus these two aspects should be carefully considered in coffee classification during coffee-bean selection, in addition to their physical aspects, such as size, color and defective beans (http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_tc_browse.htm?commid=47950S. The Cupping Protocol of the Specialty Coffee Association of America (SCCA) (http://www.scaa.org/PDF/resources/cupping-protocols.pdf) provides an international standard for cup evaluation that, besides aroma and taste, also considers kind of roasting, equipment, and cupping preparation, among other factors. Assessment of sensory attributes consists of scoring the aroma, by smelling the dry milled sample and water infusion (Steps 1 and 2) and the Flavour plus other attributes, such as aftertaste, acidity, body, and balance, by tasting the brew (Step 3).

A number of studies, some of them involving molecular sensory science, have been carried out to understand the chemistry behind the overall sensory perception given by a cup of coffee, in order to identify and define key aroma and Flavour compounds [2–6]. Different analytical platforms have been used to study coffee aroma; gas-chromatography mass spectrometry and/or olfactometry (GC-MS, GC-O) were the analytical techniques of choice. Conversely, several sampling approaches were used to extract and concentrate the Flavour components directly from the ground coffee (powder) and/or from the coffee brew, including steam distillation (SD), solvent extraction (SE), fractionation of solvent extracts, simultaneous distillation–extraction (SDE), supercritical fluid extraction (SFE), pressurized-fluid extraction, Soxhlet extraction, solvent-assisted Flavour evaporation (SAFE), microwave-assisted hydrodistillation (MAHD), headspace (HS) techniques, and solid-phase microextraction (SPME)[7]. Whatever the approach, sample preparation is still the bottle-neck of the analytical process, since it must provide a consistent and meaningful picture of the sensory-informative components. An effective sample preparation technique requires some key requisites, including (a) the possibility of tuning extraction selectivity by modifying physico-chemical characteristics of extractants and sampling conditions; (b) use of methods involving mild interactions to limit artifact formations (e.g. partition (sorption) *versus* adsorption as extraction mechanism); (c) the possibility of full automation, and of integrating the extraction step with the analytical system.

However, both compositional data and sensory information alone do not fully explain the importance of key compounds, nor indicate which of them cause distinct sensory attributes. Recently, Dunkel et al. (2014) considered more than 10,000 volatiles detected in food and determined that the specific odor code of a food is due to between 3 and 40 key odorants. Moreover, Flavour implies a multisensory process involving distinct sensory properties (mainly odors and tastes) that are closely integrated and reinforce one another [8,9,10]. These interactions may be due to different compounds that mutually influence the perceived Flavour, involving interactions between odorants (odor synesthesia) and/or odorants and tastes (chemesthesis)[11]. An important contribution to clarifying how our sense of olfaction deconvolves a complex food odor at the molecular level has been made by the genetic codification of the olfactory receptors, and the exploration of the chemistry-biology synergism of olfaction[2,12]. Very recently, Geithe et al. demonstrated that a recombined butter aroma, resulting from four odor-active compounds, each tested on *in vitro* class-I odor receptors, showed different and concentration-dependent patterns of activation [13].

Although several studies have sought to clarify the link between sensory properties and chemical composition, including through multivariate data analysis (MVA) [2,14–20], the challenge of explaining the pleasure of a coffee-experience at the molecular level still remains, mostly because of the limits of the strategies used to collect information (number and kind of samples, standardization of the samples, precision and accuracy)[21].

This study is part of a wider project exploring the correlation between the chemical composition of coffee volatile fraction and the sensory properties of the beverage; the end-goal is to develop an instrumental analysis approach complementary to human sensory profiling[8,14,15,22]. In particular the study compares chemical information related to coffee aroma and Flavour obtained with three different sampling approaches, combined in on-line or in off-line mode with GC-MS, taking the SCAA protocols for cup evaluation as reference. Because of the wide range of volatility, water solubility, and concentration of the most significant components of the coffee matrix, three different sampling approaches were tested for the reliability of characterization of the aroma and Flavour profiles, and to evaluate their compatibility with the cupping evaluation in coffee selection for quality control. Aroma evaluation (steps 1 and 2 of the SCAA cupping protocol) was associated to Headspace Solid Phase Microextraction (HS-SPME) of roasted coffee powders and the corresponding brews; aroma and taste evaluation (step 3) was combined with *in*-solution sampling of the brew by SBSE (Stir Bar Sorptive Extraction). The ability of each optimized method to discriminate and describe the investigated samples was compared by multivariate analysis, to determine whether it provided consistent and/or complementary information also in connection to the sample sensory properties defined by a trained panel according to SCAA cupping protocols.

2. MATERIALS AND METHODS

2.1 Reagents and Matrices.

Coffees samples, consisting of roasted coffee ground to suit a coffee-filter machine, were kindly supplied over a period of 9 months by Lavazza Srl (Turin, Italy).

Eight coffee samples with distinctive sensory notes, originating from different countries (Ethiopia, Papua New Guinea, Colombia, Brazil, India, Indonesia, Java, and Uganda), of the species *Coffea Arabica* L. (Arabica) and *Coffea canephora Pierre* (Robusta), were analyzed (Table 1). Each coffee origin was analyzed in five replicates; each replicate was produced by a fresh cycle of roasting and grinding, starting from the same batch of green coffee beans (n=40). The roasting degree of each sample was carefully measured by ground bean light reflectance, with a single-beam Neuhaus Neotec Color Test II instrument (Genderkesee, Germany) at a wavelength of 900 nm on 25-30g of ground coffee. Roasting degree was set at 55°Nh, in order to be close to the international standardization protocol for cupping (SCAA, 2015). Samples were roasted within 24 hours prior to cupping and left for at least 8 hours to stabilize. For clarity of exposition, samples in the text are labeled with their origins.

The coffee brew was prepared from 18g of coffee powder and 300mL of water, using a Lavazza "Xlong" coffee filter machine. Tridecane (*n*-C_{13}) in Dibuthylphtalate (DBP), used as internal standard (ISTD), were purchased from Sigma-Aldrich (Milan-Italy).

Table 5 List and characteristics of the coffee samples used in this study.

Sample acronym	*Sample Name*	*Species*	*Treatment*	*Sensorial Attribute*
BRA	BRAZIL LA2	Arabica	Natural	Nutty, quite acid, rich
COL	COLOMBIA CL1	Arabica	Washed	Flowery, Acid
JAV	JAVA WB1 MB	Robusta	Washed	Nutty
UGA	UGANDA STD	Robusta	Natural	Spicy
PNG	PAPUA NG Y	Arabica	Washed	Fruity
INDIA	INDIA ARAB CHERRY	Arabica	Natural	Astringent, quite bitter
INDO	INDONESIA EK1	Robusta	Natural	Woody, Bitter
KAFA	ETIOPIA KAFA GR. 3	Arabica	Natural	Flowery/Fruity, rather Acid

2.2. Sample preparation techniques.

HS-SPME of the coffee powder: 1.500 $\pm$ 0.010 g of powder were weighed in a septum-sealed gas vial (20mL); the resulting headspace was sampled through the PDMS/DVB SPME fiber for 40 minutes at 50°C with an agitation speed of 350rpm. The internal standard was loaded onto the fiber [23] in advance by sampling 5µL of a 1000mg/L solution of *n*-C_{13} in DBP into a 20mL headspace vial for 20 min at 50°C, agitation speed of 350rpm.

HS-SPME of the brew: a volume of 4.5mL of brew in a septum-sealed gas vial (20mL) were sampled through the SPME fiber for 40 min at 50°C with an agitation speed of 350rpm. The internal standard was loaded onto the SPME fiber in advance by sampling 5µL of a 1000mg/L *n*-C_{13} in DBP solution in a 20mL headspace vial for 20 min at 50°C, agitation speed of 350rpm [23].

SBSE of the brew: a volume of 13mL of the brew in a 20mL septum-sealed glass vial were added to 5mL of the 1mg/L *n*-C_{13} in water solution and sampled with a PDMS Twister® for 40 min at 50°C.

Brew preparation is already described in paragraph 2.1. Each sample was analyzed twice with each of the sampling methods adopted.

2.3 Standardization of sampling techniques.

SPME devices and PDMS/DVB fused silica 1 cm long fibers from the same lot were from Supelco (Bellefonte, PA, USA). Before use, all fibers were conditioned as recommended by the manufacturer, and tested to evaluate the consistency of their performance *versus* a reference roasted coffee sample [24]. Normalized peak areas collected from the entire set of analyses (three replicates per sample) and from all fibers (n=9) were submitted to analysis of variance (ANOVA). Only fibers that do not showed statistical differences through the one-way ANOVA test (confidence interval 95%). The same protocol was applied to SBSE devices (1cm x 0.5mm PDMS coated Twister®, Gerstel GmbH & Co. KG).

2.4 Analysis Conditions.

HS-SPME analysis was carried out with a QP2010 GC-MS system (Shimadzu - Milan, Italy) equipped with an autosampler combi-PAL AOC 5000 Autoinjector (Shimadzu - Milan, Italy).

SBSE sampled analytes were thermally desorbed from the Twisters® using a thermal desorption system (TDS-2; Gerstel, Mülheim, Germany) installed on an Agilent 6890plus gas chromatograph coupled with a MSD Agilent 5973D. A cooled injection system (CIS-4PTV; Gerstel, Mülheim, Germany) was used to focus the thermally desorbed analytes cryogenically at -50 °C with liquid carbon dioxide.

HS-SPME-GC-MS chromatographic conditions: injector temperature: 230°C; injection mode, splitless; carrier gas, helium (2mL/min); fiber desorption time and reconditioning, 5min; column, SGE SolGelwax (100% polyethylene glycol) 30 m x 0.25 mm d_c x 0.25 µm d_f (SGE- Melbourne, Australia); temperature program, from 40°C (1min) to 200°C at 3°C/min, then to 250°C (5min) at 10°C/min. *MS conditions:* ionization mode: EI (70eV); scan range: 35-350 amu; ion source temperature: 200°C; transfer line temperature: 250°C.
SBSE-GC-MS chromatographic conditions: injector temperature: 250°C; injection mode, splitless; carrier gas, helium (1mL/min); column, SGE SolGelwax (100% polyethylene glycol) 30 m x 0.25 mm d_c x 0.25 µm d_f (SGE- Melbourne, Australia); temperature program, from -30°C (0min) to 40°C (1min) at 60°C/min, then to 200°C (0min) at 3°C/min, then to 250°C (5 min) at 10°C/min.
MS conditions: ionization mode: EI (70eV); scan range: 35-350 amu; ion source temperature: 230°C; transfer line temperature: 280°C.
TDS temperature program: from 30°C to 250°C at 60°C/min; hold time at final temperature: 10min; delay time: 0min; initial time: 1 min.
CIS temperature program: from -50°C to 250°C at 12°C/s; hold time at final temperature: 5min; equilibration time: 0.1min; initial time: 0 min.

2.5 Identification of Volatile Components.
Aroma compounds sampled from headspace of powder and from brew were identified by comparing their calculated linear retention indices and their mass spectra to those of authentic samples or, tentatively, to those collected in homemade or commercial libraries (Wiley *7N* and Nist 05 ver 2.0 Mass Spectral Data) or reported in the literature.

2.6 Sensory analysis
The fourty samples were submitted to a sensory evaluation by a panel of five experts using 18 g of roasted and ground coffee in 300 mL of hot water according to the SCAA protocols (SCAA, 2014). The protocol implies three tasting steps after roasting to a fixed color (55-60° Nh) and eight hours of sample stabilization: i) evaluation of the aroma by sniffing the dry grounded coffee, ii) evaluation of the aroma by sniffing the brew three minutes after its preparation and stirring, and iii) 8-10 minutes after Flavour evaluation. Other attributes such as aftertaste, acidity, body, and balance are evaluated by tasting the brew by spraying it in the mouth to maximize retro-nasal vapors. The cup quality was assessed for several attributes, among them this study considered: Flavour (floral, fruity, woody, nutty, spicy), acidity, bitterness, body (mouthfeel), astringency, and overall quality. The quality and intensity of each attribute were evaluated simultaneously by using a scale varying from 1 to 10.

2.7 Data processing.
Data were collected with a Shimadzu GCMS Solution 2.5SU1, and an Agilent ChemStation D.02.00.275. Principal Component Analysis (PCA) was used to visualize sample groups and to compare information provided by each sampling. PCA based on Pearson correlation coefficient was carried out on normalized ISTD data. Statistical analysis one-way ANOVA and PCA were done by XLSTAT (version 2015.5.01.23164) copyright Addinsoft 1995-2015. non-polar

3. RESULTS AND DISCUSSION

The objective evaluation of coffee quality, by correlating chemical analysis and sensory properties, requires an analytical platform that provides information appropriate to describing the human sensory experience. Coffee powder and brew, evaluated through SCAA protocols, were thus analyzed with three different sampling methods, each combined with GC-MS; this resulted in chemical information describing the coffee aroma and Flavour that was in line with that employed for cup evaluation. In the following, for short, the analytical platform will be identified by the sampling used, its on-line or off-line combination with GC-MS being implicit.

3.1 Samplings comparison

A total of 117 compounds were identified (or tentatively identified) (20 compounds were unknown or not identified unequivocally) with the above platforms. Table 2 reports the list of the compounds identified with each sampling with their Linear Retention Indices (I^Ts). The highest number of compounds (96) were identified in the headspace of the coffee powder, followed by HS-SPME (72) and SBSE (53) of the brew.

Table 2 List of identified and *tentatively identified compounds in all sampling methods. (http://webbook.nist.gov/chemistry/name-ser.html)

#	Compound Name	Calc. I^Ts	Lit. I^Ts	HS-SPME powder	HS-SPME brew	SBSE brew
1	*Acetaldehyde*	*706*	*723*	*X*	*X*	
2	*Acetone*	824	835	X	X	
3	*Methyl acetate*	828	839	X	X	
4	*Furan, 2-methyl-*	885	864	X	X	
5	*2-Butanone*	906	905	X	X	
6	*Butanal, 2-methyl-*	914	931	X	X	
7	*Butanal, 3-methyl-*	918	936	X	X	
8	*Furan, 2,5-dimethyl-*	950	939	X	X	
9	*2,3-Butanedione*	978	963	X	X	
10	*2,3-Pentanedione*	1058	-	X	X	
11	*2-Vinylfuran*	1071	1085	X	X	
12	*Hexanal*	1080	1098	X		
13	*2,3-Hexanedione*	1130	1110	X		
14	*1H-Pyrrole, 1-methyl-*	1137	1140	X	X	X
15	*2-Vinyl-5-methylfuran*	1151	1152	X		X
16	*Pyridine*	1177	1177	X	X	X
17	*Pyrazine*	1209	1206	X	X	
18	*Furfuryl methyl ether*	1238	-			X
19	*pyrazine, methyl-*	1262	1268	X	X	X
20	*2-butanone, 3-hydroxy-*	1281	1285	X		
21	*pyridine, 3-methyl-*	1291	1297			X
22	*2-propanone, 1-hydroxy-*	1297	1318	X		
23	*pyrazine, 2,5-dimethyl-*	1317	1321	X	X	X
24	*pyrazine, 2,6-dimethyl-*	1324	1327	X	X	X
25	*pyrazine, ethyl-*	1329	1343	X	X	X

#	Compound Name	Calc. I^Ts	Lit. I^Ts	HS-SPME powder	HS-SPME brew	SBSE brew
26	pyrazine, 2,3-dimethyl-	1341	1354	X	X	X
27	1-hydroxy-2-butanone	1370	1381	X		
28	pyridine, 3-ethyl-	1374	1384	X	X	X
29	pyrazine, 2-ethyl-6-methyl-	1382	1392			X
30	pyrazine, 2-ethyl-5-methyl-	1386	1399	X	X	X
31	pyrazine, 2-ethyl-3-methyl- + pyrazyne, trimethyl	1399	1400	X	X	X
32	unknown 1	1405	-	X	X	
33	2-n-propylpyrazine	1413	1425	X	X	X
34	unknown 2	1418	-	X	X	
35	2-furanmethanethiol	1432	1440	X		
36	2-ethyl-3,6-dimethylpyrazine	1441	1449	X	X	X
37	acetic acid	1446	1454	X	X	X
38	pyrazine, 2,3-diethyl-	1451	1463			X
39	pyrazine, 2,6-diethyl-	1457	1458	X	X	X
40	furfural	1462	1467	X	X	X
41	acetoxyacetone	1468	1467	X	X	X
42	furfuryl methyl sulphide	1483	1496	X	X	X
43	pyrazine, 2-methyl-6-vinyl-	1483	1491			X
44	pyrazine, 3,5-diethyl-2-methyl-	1489	1491	X	X	X
45	3(2h)-furanone, 2,5-dimethyl-	1496	1490	X		
46	furfurylformate	1497	1497	X		
47	2-acetylfuran	1500	1498	X	X	X
48	2-cyclopenten-1-one 3 methyl + 3,5-diethyl-2-methylpyrazine	1509	1509	X		
49	benzaldehyde	1512	1522			X
50	1h-pyrrole	1513	1525	X	X	
51	1-(2-furyl)-2-propanone	1519	1519			X
52	belzhaldehyde + 2-methyl-3(2h) tiophenone + furan-2-yl-propan-2-ol	1519	1520	X	X	
53	2 methoxy 3 isobutyl pyrazine	1523	1531			X
54	2-oxopropylpropanoate	1531	1531	X	X	
55	furfuryl acetate	1538	1539	X	X	X
56	unknown 4	1564	-	X		
57	2-furancarboxaldehyde, 5-methyl-	1567	1558			X
58	5 methyl furfural	1570	1562	X	X	
59	2,3-butandiolo + 1-(5-methyl-2-furyl)2-propanone	1575	1582	X		
60	pyrazine, (1-methylethenyl)-	1590	-	X		
61	2-furanmethanol, propanoate	1598	1603	X	X	X
62	furan, 2,2'-methylenebis-	1606	1606			X
63	2-furfurylfuran	1608	-	X	X	
64	(5h)-5-methyl-6,7-dihydrocyclopentapyrazine	1611	1611	X		
65	1h-pyrrole-2-carboxaldehyde, 1-methyl-	1614	1635	X	X	X
66	butanoic acid, 4-hydroxy-	1621	-	X	X	

#	Compound Name	Calc. I^Ts	Lit. I^Ts	HS-SPME powder	HS-SPME brew	SBSE brew
67	unknown 6	1630	-	X	X	
68	2-isopropenylpyrazine	1633	1633	X	X	
69	1-(2-furyl)-3-butanone	1641	1642			X
70	2,5-dihydro-3,5-dimethyl-2-furanone	1642	1640	X		
71	2-acetyl-1-methylpyrrole	1646	1667			X
72	ethanone, 1-(1-methyl-1h-pyrrol-2-yl)-+2-acetyl-5-methyl pyrrole	1649	-	X	X	
73	furfuryl alcohol	1661	1664	X	X	X
74	3-methylbutanoic acid	1667	1670	X	X	
75	furan, 2-(2-furanylmethyl)-5-methyl-	1677	-			X
76	2-furfuryl-5-methylfurane	1678	-	X	X	
77	2-acetyl-3-methylpyrazine*	1686	1636	X	X	
78	furfurylpentanoate + others unknown compounds	1694	1719	X		
79	pyrazine, 2-methyl-5-(1-propenyl)-	1702	1702	X	X	
80	3-methyl-1,2-cicloheanedione	1708	1719	X	X	
81	unknown 11	1709	-	X		
82	1-acetyl-1,4-dihydropyridine*	1716	-	X	X	
83	unknown 12	1726	-	X		
84	unknown 13	1729	-	X	X	
85	unknown 13b	1734	-	X	X	
86	unknown 14	1745	-	X		
87	unknown 15	1750	-	X	X	
88	methyl salicylate	1758	1754			X
89	methyl nicotinate + others unknown compounds	1767	1778	X		
90	unknown 17	1772	-	X		
91	unknown 18	1777	-	X	X	
92	unknown 19	1786	-	X	X	
93	unknown 20	1791	-	X		
94	beta-damascenone	1810	1816			X
95	n-furfurylpyrrole	1820	1839	X	X	X
96	guaiacol	1853	1866	X	X	X
97	unknown 21	1860	-	X	X	
98	2-cyclopenten-1-one, 2-hydroxy-3-methyl-	1885	-	X	X	
99	2-cyclopenten-1-one, 3-ethyl-2-hydroxy-	1885	-	X	X	
100	unknown 21b	1893	-	X	X	
101	trans-furfurylideneacetone	1897	-	X	X	
102	benzeneethanol	1902	1912	X		
103	maltol	1952	1960	X	X	X
104	2-acetylpyrrole	1962	1971	X	X	X
105	difurfuryl ether	1980	1977	X	X	X
106	phenol	1996	2013			X
107	unknown 23	1997	-	X	X	
108	1h-pyrrole-2-carboxaldehyde	2012	2035	X	X	X

#	Compound Name	Calc. I^Ts	Lit. I^Ts	HS-SPME powder	HS-SPME brew	SBSE brew
109	*4-ethyl-guaiacol*	2021	2037	X	X	X
111	*unknown 23b*	2073	-	X	X	
112	*4-vinyl-guaiacol*	2185	2193	X	X	X
113	*1-furfuryl-2-formyl pyrrole*	2230	2234			X
114	*1h-indole*	2411	2443			X
115	*benzophenone*	2443	-			X
116	*difurfuryldisulfide*	2536	2536	X		X
117	*caffeine*	2838	-			X

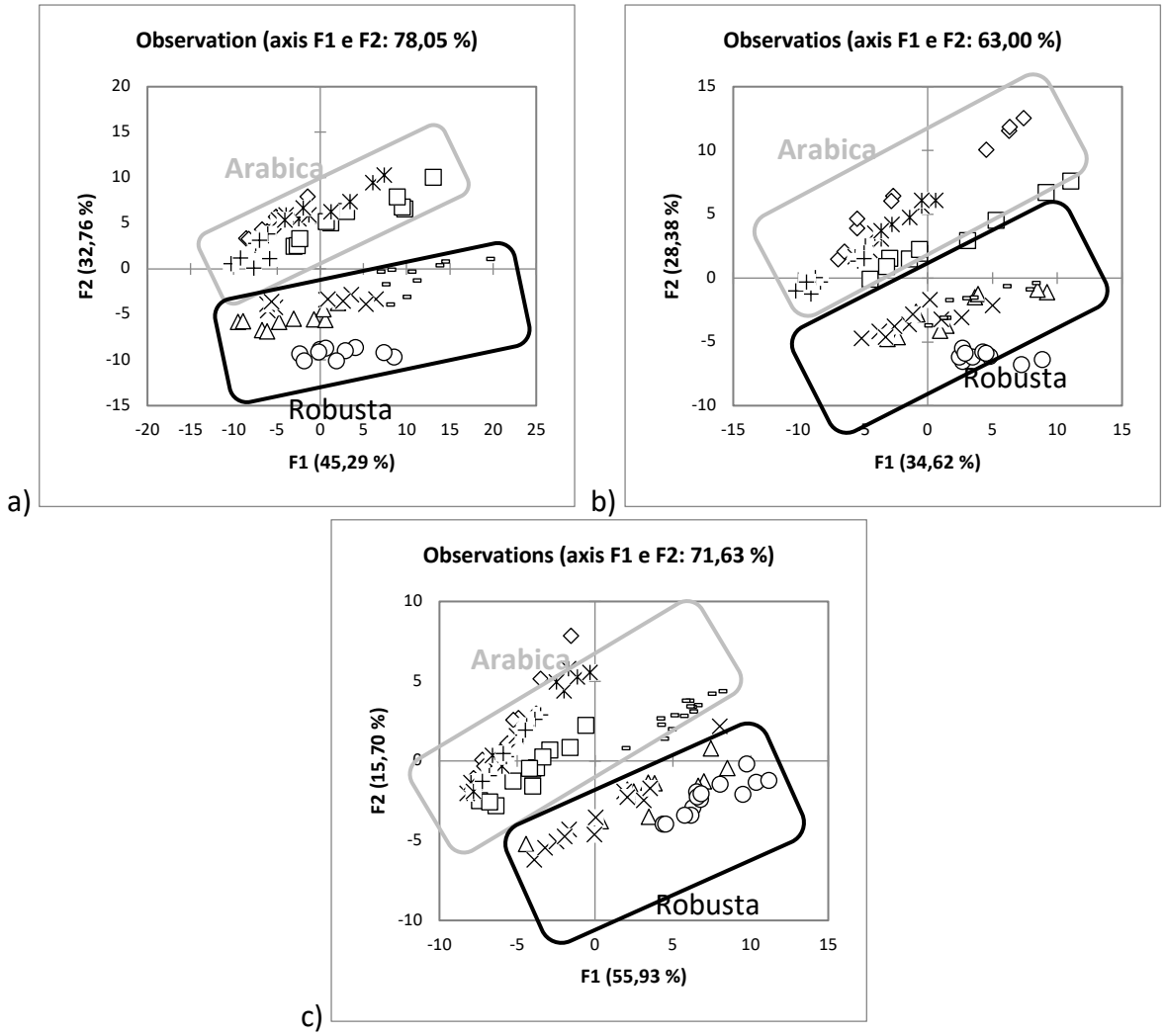


Figure 1 PCA score plots of a) HS-SPME of the coffee powder; b) HS-SPME of the brew; c) SBSE of the brew. Autoscale pre-processing. Legend: BRA: □; COL: ◊; JAV: Δ; UGA: X; PNG: *; INDIA: -; INDO: ◯; KAFA: +

The chemometric approach (PCA) was used to obtain as much information as possible from the three sampling methods: each sample (observations) is described by different compounds (variables), with their own analytical response. Figure 1 reports the PCA score plots of a) HS-SPME of coffee powder, b) HS-SPME of coffee brew, and c) SBSE of coffee brew. The comparison of the PCA results from the brews sampled by HS-SPME (b) and SBSE (c) shows a similar distribution of

the samples on the score plot. Similar discrimination of samples is also obtained by the HS-SPME of the powder (a); this means that independently of the sampling approach applied, the information derived from the chemical profiles of the samples is the same, as it is also evident from the total explained variance obtained with PCA elaborations. Two large groups were recognizable along the PC2 that, as expected, were chiefly characterized by species, i.e. Arabica or Robusta. INDIA samples were the only exception, being close to Robusta samples although classified as Arabica.

Analysis of Robusta sample profiles showed that specimens from Indonesia (INDO) can clearly be discriminated from the two other origins (JAV and UGA) on the first two PCs (Figure 2). PCA analysis on Arabica samples showed similar distribution for the three different sampling approaches (Figure 3).

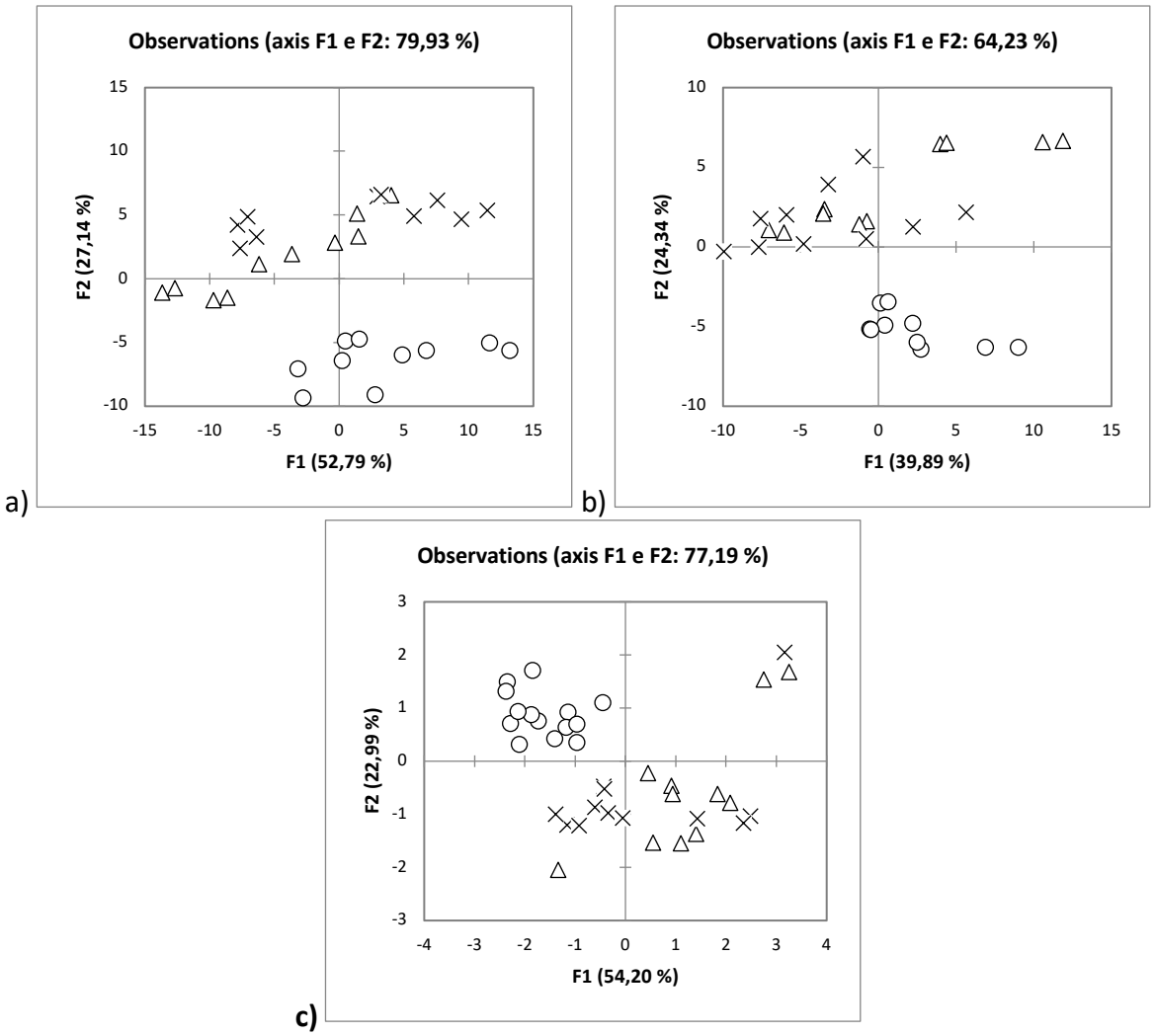


Figure 2 Robusta PCA score plots: a) HS-SPME of the coffee powder b) HS-SPME of the brews; c) SBSE of the brews. Autoscale pre-processing Legend: JAV: Δ; UGA: X; INDO: ◯

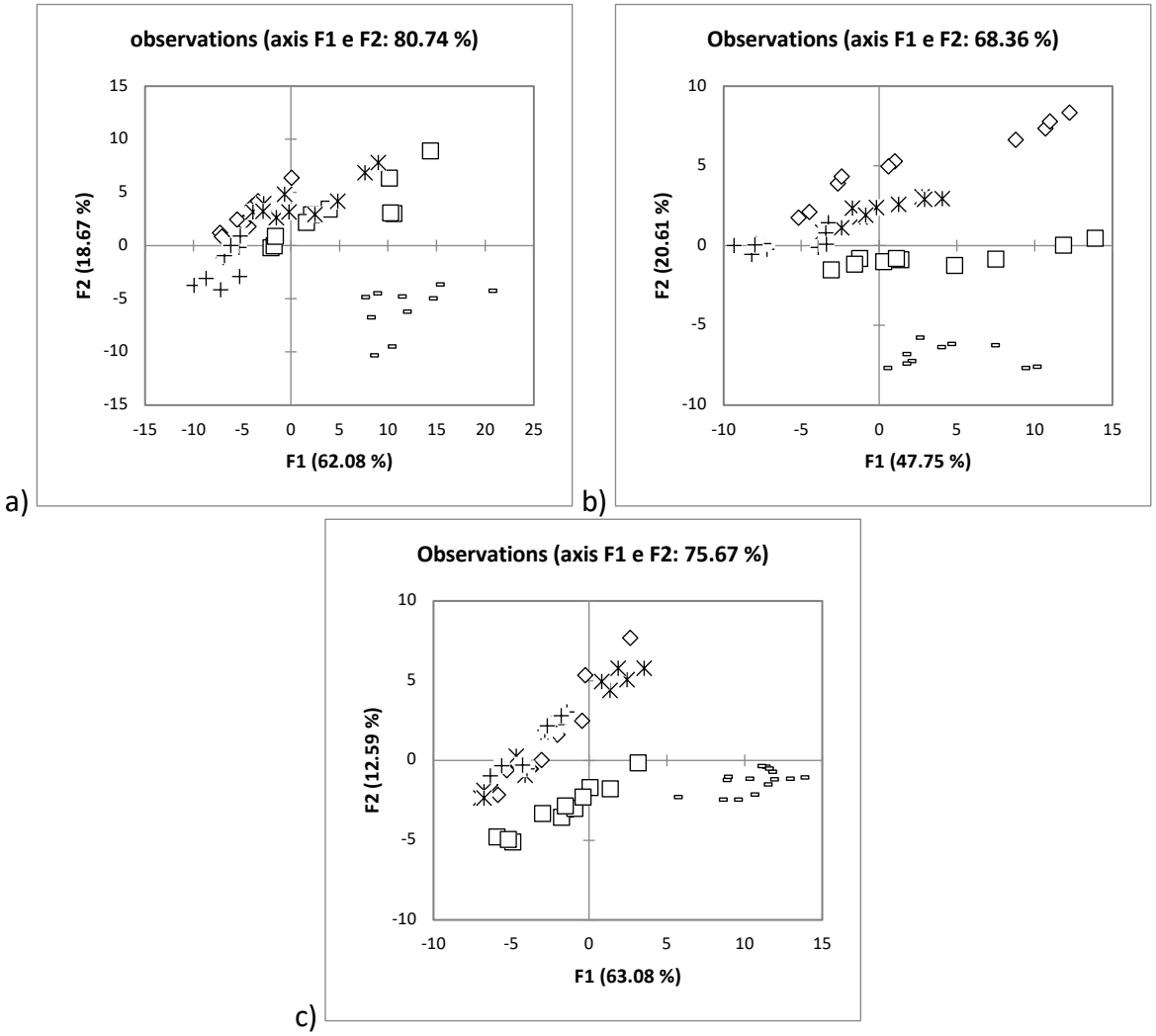


Figure 3 Arabica PCA score plots: a) HS-SPME of the coffee powder b) HS-SPME of the brews; c) SBSE of the brews. Autoscale pre-processing Legend: : BRA: □; COL : ◊; PNG: *; INDIA: -; KAFA: +

3.2 Investigation on discriminant aroma compounds with the different sampling approaches

The volatiles directly responsible for discrimination of the Robusta samples deriving from the vector projections of the original variables on PC1 and PC2 (variable $\cos^2$) are listed in Table 3, together with their odor description. For the sake of clarity, these components will henceforth be indicated as Direct Discriminant Compounds (DDCs). PCA determined different DDCs for each sampling method, partly because the methods are based on different principles, employ different sampling materials (PDMS/DVB SPME fibers for headspace, and PDMS Twisters® for *in*-solution sampling), and are applied to different matrices (coffee powder and brew) (Table 3). Further, compounds directly responsible for sample discrimination in SBSE sampling of the brew, which may be considered the most representative sampling technique for Flavour evaluation, cannot be the same as those for HS-SPME sampling of the coffee powder, because the intrinsic physical-chemical properties of those compounds influence their recovery. The relationship between the role of each compound in sample discrimination and their physico-chemical properties (EPI Suite v3.10 developed by the EPA's Office of Pollution Prevention Toxics NS Syracuse Research Corporation (SRS) 2000 U.S.) was thus studied, to investigate in greater depth why different compounds may play the same roles in sample discrimination, independently of the technique adopted. Most of the DDCs with SBSE of the brew are slightly soluble in water and relatively non

polar, i.e. with high $k_{o/w}$ (Table 3). Conversely, DDCs in the HS-SPME volatile fraction of the coffee powder generally present high volatility (expressed as Vapour Pressure, VP) and low $k_{o/w}$ (below 1) (Table 3). Similarly to SBSE, HS-SPME of the brew includes compounds extracted during brewing whose relatively high water solubility has less influence on the composition of the headspace, since they are retained in the aqueous phase [25,26]. Moreover, the coffee powder may be considered a fatty matrix, and thus polarity may also influence migration into the headspace, and non polar compounds (high $k_{o/w}$ values) may undergo a more severe matrix effect.

Table 3 DDCs extracted from processing Robusta samples. Brews sampled by SBSE or HS-SPME and HS-SPME of the powder, with their relative odor descriptors and physico-chemical properties. Letters near the name indicate the sampling approaches where each compound was recovered: SBSE: A; HS-SPME pow: B; HS-SPME brew: C. * The Good Scents Company, [&] http://www.iso.org, [+] Blank et al.

Compound Name	***Odour Description**[*,&,+]*	***Water solubility (mg/L)***	***Log $K_{O/W}$***	***VP (mm Hg at 25 °C)***	***Henrys LC (VP/Wsol) (atm-m3/mole)***
1-acetyl-1,4-dihydropyridine ***(C)***	-	-	-	-	-
*1H-Pyrrole-2-carboxaldehyde (**A; B; C)***	Musty	3.43E+04	0.6	0.09	3.13E-07
1-Hydroxy-2-butanone ***(B)***	Sweet coffee musty grain malt butterscotch	7.21E+05	-0.29	0.77	1.24E-07
2,3-Butanedione ***(B)***	Buttery	2.00E+05	-1.34	56.8	7.95E-06
2,3-Pentanedione ***(B; C)***	Buttery	6.16E+05	-0.85	31.1	6.65E-06
2-Butanone, 3-hydroxy- ***(B)***	Buttery	8.33E+05	-0.36	2	2.78E-07
2-cyclopenten-1-one, 2-hydroxy-3-methyl- ***(C)***	Caramellic-spicy, maple-like	8.50E+03	1.29	0	6.68E-08
2-Furancarboxaldehyde, 5-methyl- ***(A)***	Caramel	2.91E+04	0.67	1.38	
2-furfuryl-5-methylfurane ***(B)***	-	6.40E+01	1.96	2.89	1.96E-04
2-Furfurylfuran ***(B; C)***	Roast	2.14E+02	2.99	0.26	2.36E-04
2-Oxopropylpropanoate ***(B)***	-	1.10E+04	1.2	31.5	4.02E-04
2-Propanone, 1-hydroxy- ***(B)***	Caramel	7.44E+01	-0.78	1.74	1.70E-07
2-Vinyl-5-methylfuran ***(B; C)***	-	2.21E+03	1.96	2.89	1.96E-04
3(2H)-Furanone, 2,5-dimethyl- ***(B)***	Fruity, caramellic	4.63E+04	0.43	1.66	5.29E-06
4-Ethylguaiacol ***(A)***	Spicy	6.94E+02	2.38	0.02	
5 Methyl Furfural ***(B;C)***	Caramel	2.91E+04	0.67	0.69	3.41E-06
Acetaldehyde ***(B)***	Pungent ethereal aldehydic fruity	2.57E+05	-0.34	910	1.72E-04
Acetic acid ***(B)***	sharp pungent sour vinegar	4.76E+05	-0.17	15.7	2.86E-06
Acetoxyacetone ***(A; B; C)***	Fruity	1.52E+05	-0.19	1.49	1.50E-06
Benzaldehyde ***(A)***	Strong sharp sweet bitter almond cherry	6.10E+03	1.71	1.01	
Butanal, 3-methyl- ***(C)***	Aldehydic	1.12E+04	1.23	51.6	5.21E-04
Difurfuryl ether ***(C)***	Coffee, nutty, earthy	7.11E+02	2.22	0.02	7.48E-06
Furan, 2-(2-furanylmethyl)-5-methyl- ***(A)***	Hearthy, mushroom	6.41E+01	3.53	0.07	

Compound Name	***Odour Description**,&,+***	***Water solubility (mg/L)***	***Log $K_{O/W}$***	***VP (mm Hg at 25 °C)***	***Henrys LC (VP/Wsol) (atm-m3/mole)***
Furan, 2,2'-methylenebis- ***(A)***	Roast	2.17E+02	2.99	0.26	
Furfural ***(A, B; C)***	sweet woody almond fragrant baked bread	5.36E+04	0.83	2.32	5.48E-06
Furfuryl methyl sulphide ***(A)***	Vegetable	1.84E+03	2	1.37	
Guaiacol ***(C)***	Spicy	2.09E+03	1.88	0.06	5.16E-06
4-ethyl-guaiacol ***(C)***	Spicy	6.94E+02	2.38	0.02	7.16E-06
4-vinyl-guaiacol ***(C)***	Woody	9.26E+02	2.24	0.01	1.64E-06
Hexanal ***(B)***	fresh green fatty aldehydic grass leafy fruity sweaty	3.52E+03	1.78	9.57	3.58E-04
Pyridine, 3-ethyl- ***(A; B; C)***	Tobacco	8.48E+04	1.84	2.53	3.29E-06

These considerations are clearly explained by the comparison of normalized percent areas of some DDCs obtained with the three sampling approaches. 3-Ethyl pyridine and furfural (i.e. two DDCs with similar physico-chemical characteristics) are differently recovered by SBSE, 3-ethyl pyridine predominating because of its higher $k_{o/w}$, while furfural, being more polar, is less retained by the fatty matrix and more easily released into the headspace. Conversely, by comparing HS-SPME of the brew to SBSE, the more polar furfural is less recovered than does 3-ethyl pyridine from the headspace of the brew and recovered to a greater extent by SBSE (Figure 4). Acetoxyacetone is highly concentrated in the headspace of coffee powder and is recovered better by SBSE than by HS-SPME from the brew, because of its high solubility in water. 1-H-Pyrrole-2-carboxaldehyde contributes similarly to HS-SPME from coffee powder and brew, but having a medium-low $k_{o/w}$, good water solubility and low VP, its accumulation in the headspace is limited.

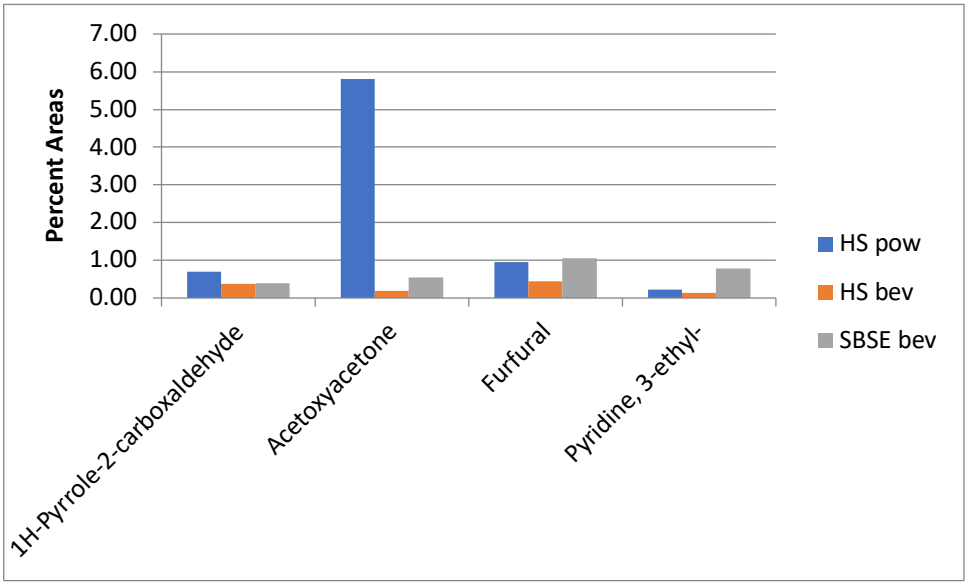


Figure 4 Comparison between normalized percentage contributions of the common direct discriminant compounds in the three sampling approaches under study.

Moreover, DDCs from SBSE can also be correlated to other compounds from the HS-SPME sampling, "indirect markers" or CDCs (Correlative Discriminant Compounds), which are indirectly involved in the discrimination of the coffee powder by HS-SPME. CDCs can be defined through the Pearson correlation coefficient (r), used here to assess the degree of linear association between variables (peak area vectors) defined by the different samplings, r values > 0.8 were taken as cut-off point. From the chemometric standpoint, variables with high r values with DDCs, within the PCA elaboration of the HS-SPME of coffee powder, are redundant for the purpose of explaining sample behavior with this approach. Therefore, some of them may be eliminated without lacking in quality of discrimination, because they are dependent variables and provide the same information of DDCs, in terms of sample definition.

Table 4 Compounds present in HS-SPME of the powder that are closely correlated with DDCs of SBSE. The DDCs in common between the two sampling techniques are in bold type. Compounds with a direct discriminant role in SBSE or HS-SPME of coffee powder are marked with an X; indirect markers (CDCs) are in italics.

Compounds	***DDCs in SBSE of the brew***	***DDCs in HS-SPME of the powder***
1-Hydroxy-2-butanone		X
1H-Pyrrole-2-carboxaldehyde	**X**	**X**
1H-Pyrrole-2-carboxaldehyde, 1-methyl-		
2-acetylpyrrole		
2-butanone		
2-Butanone, 3-hydroxy-		X
2-oxopropylpropanoate		X
2-Cyclopenten-1-one, 2-hydroxy-3-methyl-		
2-furfurylfuran		X
2-n-propylpyrazine		
2-Propanone, 1-hydroxy-		X
2-Vinyl-5-methylfuran		X
2,3-butanedione		X
2,3-pentanedione		X
2-cyclopenten-1-one 3 methyl+ 3,5-diethyl-2-methylpyrazine		
3(2H)-Furanone, 2,5-dimethyl-		X
5 methyl furfural		X
Acetic acid		X
Acetone		
Acetoxyacetone	**X**	**X**
Acetylfuran		
Ethanone, 1-(1-methyl-1H-pyrrol-2-yl)- + 2-acetyl-5-methyl pyrrole		
Furan, 2-methyl-		
Furfural	**X**	**X**
Furfuryl alcohol		
Furfurylformate		
Furfuryl methyl sulphide	X	
Guaiacol		
4-ethyl-guaiacol	X	

Compounds	***DDCs in SBSE of the brew***	***DDCs in HS-SPME of the powder***
Hexanal		X
Methyl acetate		
Pyrazine, (1-methylethenyl)-		
Pyrazine, 2-ethyl-3-methyl- + Pyrazyne, trimethyl		
Pyrazine, 2-methyl-6-(1-propenyl)-		
Pyrazine, 2,3-dimethyl-		
Pyrazine, 2,6-diethyl-		
Pyrazine, 3,5-diethyl-2-methyl-		
Pyridine		
Pyridine, 3-ethyl-	**X**	**X**
Unknown 1		X
Furfurylpentanoate + other unknown compounds		
Unknown 12		X
Unknown 13		X
Unknown 14		X
Unknown 17		X
Unknown 2		X
Unknown 21		X
Difurfuryl ether		
Unknown 23b		
(5h)-5-methyl-6,7-dihydrocyclopentapyrazine		
Unknown 6		
2-isopropenylpyrazine		
2,5-dihydro-3,5-dimethyl-2-furanone		
2-furfuryl-5-methylfurane		

The consistency between the three samplings was confirmed by including DDCs of the SBSE in the data correlation matrix of the HS-SPME; resulting in a close correlation with 56 compounds identified in the HS-SPME of the powder. Twenty-four of them were also HS-SPME DDCs, while the remaining 32 were CDCs of this method. Table 4) reports the compounds identified in the HS-SPME-GC-MS profile of the coffee powder having high r (> 0.8) with SBSE DDCs. This means that DDCs from *in*-solution SBSE sampling, direct (DDCs) or indirect markers (CDCs) of the HS-SPME of the coffee powder, provide chemical information for sample differentiation that is related to the sample different chemical processing and sensory characteristics, and, as a consequence, to their chemical pathways of formation. In other words, a compound that is highly soluble in water may not play a direct role in the discrimination of coffee powder headspace but, thanks to its solubility, it may be solubilized during brewing in large amounts, and thus play an important role in the discrimination of beverages. Conversely, a CDC may have different physico-chemical properties but provides the same kind of chemical information as a DDC in the discrimination of samples with different sensory characteristics**.** Similar observations can be made for the role played by SBSE DDCs in samples discrimination obtained by the HS-SPME of the brews (Table 5). These considerations resulted also valid for the analysis of INDIA Arabica samples (data not reported).

Table 5 Compounds present in HS-SPME of the brew that are closely correlated with DDC of SBSE. The DDCs in common between the two sampling approaches are in bold type. Compounds with a direct discriminant role in SBSE or in HS-SPME of the brew are marked with an X; HS-SPME brew indirect markers (CDCs) are in italics.

Compounds	***DDCs in SBSE***	***DDCs in HS-SPME brew***
1-Acetyl-1,4-dihydropyridine		X
1H-Pyrrole-2-carboxaldehyde	**X**	**X**
1H-Pyrrole-2-carboxaldehyde, 1-methyl-		
unknown 23		
2-furfurylfuran		X
2,3-pentanedione		X
5 methyl furfural		X
Acetoxyacetone	**X**	**X**
Furfural	**X**	**X**
Furfuryl methyl sulfide	X	
Guaiacol		X
4-ethyl-guaiacol	**X**	**X**
4-vinyl-guaiacol		X
Pyrazine, 2,3-dimethyl-		
Pyridine		
Pyridine, 3-ethyl-	**X**	**X**
Unknown 2		X
Unknown 21		X
Unknown 21b		
Difurfuryl ether		X
2-furfuryl-5-methylfurane		X

The similarity of the sample discrimination achieved by the three sampling approaches indicates not only that they provide complementary data, but also that they may be used interchangeably to discriminate the chemical profiles of a set of samples and can thus be applied to the problem under study. This can be explained in two complementary ways: a) the first is related to the physico-chemical properties of the components referred to as DDCs, depending on the sampling approach under study; b) the second is due to the (r) value, which correlates compounds indicative of the same change(s) in sample discrimination, and, as a consequence, of a common chemical pathway of formation.

This correlation is also clear from the chemical standpoint, if the behavior of groups of compounds of different nature (e.g. guaiacoles, pyridines, pyrazines and furans) is examined. The statistical analysis shows that these compounds are in all cases correlated with one another, irrespective of the sampling used. The comparison of data from the three approaches shows that different classes of compounds change as one, moving in the same direction, and that they always play a role in sample discrimination, irrespective of which component(s) is involved in the discrimination of a specific sampling. The formation pathways of these groups of components are induced by roasting, but also depend on the processing of the green beans. Pyrazines (generally having nutty, earthy, roasted, and green aromas) and pyridines (fishy note), principally arise from the Maillard reaction of amino acids and sugars, direct pyrolysis of amino acids and degradation of trigonelline. The roasting pathways for guaiacoles (spicy notes), for example, involve the decarboxylation of phenolic carboxylic acids and the thermal degradation of lignin; however, their formation (or concentration) in coffee aroma also depends on bacterial, fungal, and yeast enzymes, and on

glyosidic reactions occurring in the green beans [2,5]. Furans, responsible for malty, caramel, and sweet-roast notes, are formed during the roasting process through the Maillard reaction of carbohydrates, thermal oxidation of lipids, and degradation of thiamine. The discriminant furanic compounds differ with the different sampling methods but are in any case involved in the discrimination of INDO samples within Robusta, and INDIA samples for Arabica.

3.3 Relationship between chemical results and sensory cupping data

A Lavazza-trained panel determined the sensory description of the set of investigated coffee samples. The panel considered the following sensory characteristics: acid, bitter, aromatic intensity, floral, fruity, woody, nutty, spicy, together with body and astringency. Each sensory attribute was classified by the panelists on a scale from 0 to 10, where 0 signified no attribute and 10 a strong sensory attribute. Figure 5 reports the PCA scores (top) and loading plots (bottom) of the sensory evaluation of the Robusta (left) and Arabica (right) samples. Within the Robusta set, INDO samples were characterized by woody, spicy, and bitter notes; JAVA samples were slightly acid and nutty, and INDO and UGA samples were more spicy and aromatic than those from JAVA. In the Arabica set, INDIA samples were markedly woody and spicy, similarly to Robusta INDO, and presented a bitter note and strong body. BRA samples were astringent and nutty, while KAFA were the most fruity samples, also characterized by stronger acid and floral notes, followed by COL and PNG.

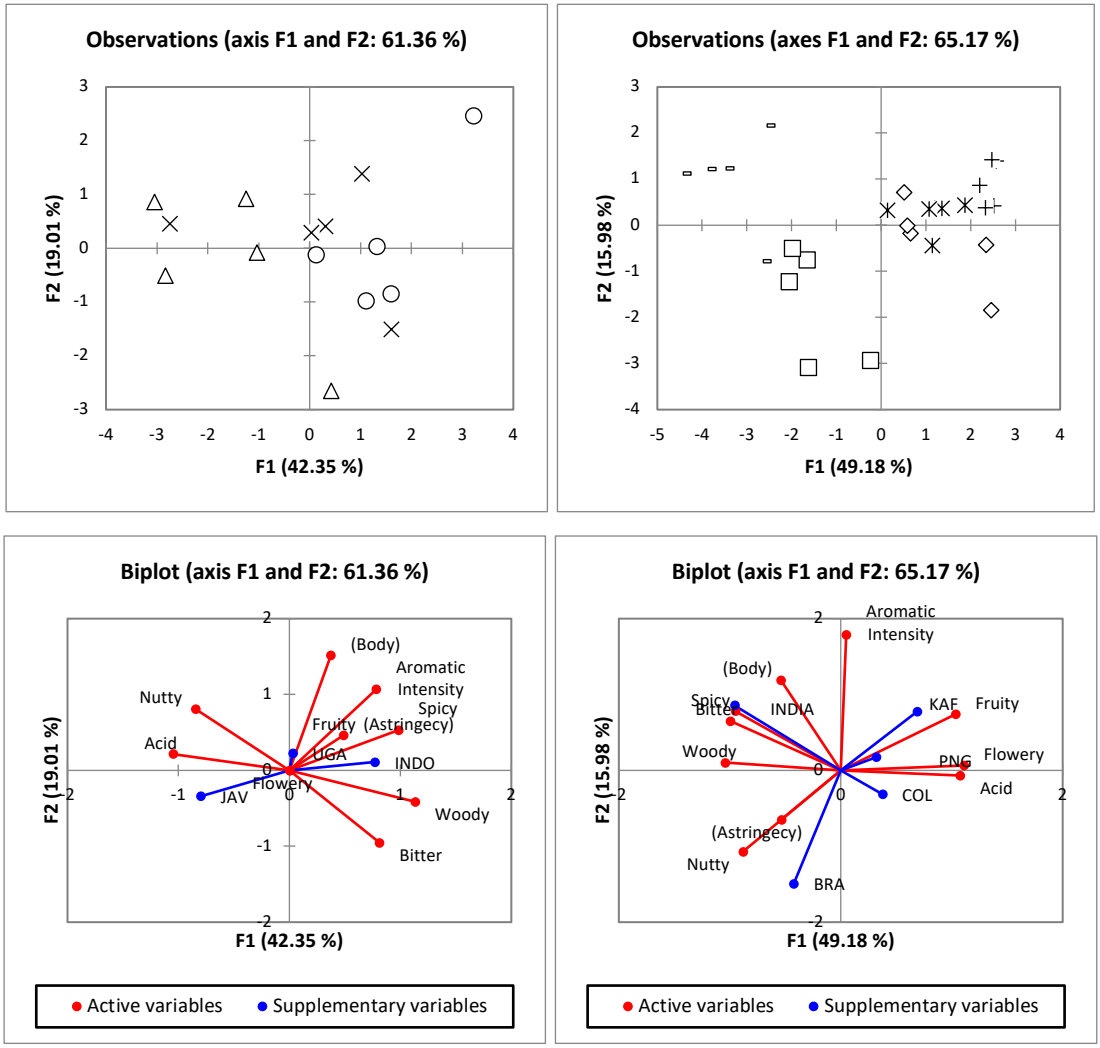


Figure 3 PCA scores (top) and loading plots (bottom) of the sensorial evaluation of the Robusta (left) and Arabica (right) samples Legend: BRA: □; COL: ◊; JAV: Δ; UGA: X; PNG: *; INDIA: -; INDO: ◯; KAFA: +

Most DDCs resulting from the chemical investigation in the different sampling approaches are known to be connected with these notes. In a chemometric investigation on Arabica samples, Ribeiro et al. showed that several compounds can be responsible for more than one sensory attribute. For instance, 3-ethyl pyridine may be responsible for acidity, Flavour, and bitterness, or 4-vinyl guaiacol for Flavour and body. However, when considered as such, their sensory attributes are not always associable to the above characteristics[18]. In particular, DDCs from the chemometric analysis of INDO and INDIA respectively for Robusta and Arabica samples include components with sensory notes that can be related directly to the sensory characteristics highlighted for these samples (Table 1 and 3). However, the peculiar odor and Flavour of these samples are not only related to the presence or absence of some compounds, but also closely depend on their relative concentrations and odor thresholds, which together are responsible for their synergistic or antagonistic effect at the receptorial level, in eliciting the sensory experience. All sampling approaches, even if with different DDCs related together to the sampling peculiarity and compound physico-chemical characteristics, are coherent with the discrimination obtained with sensory evaluation. However, the direct HS-SPME sampling of the powder requires a limited sample manipulation since it does not include the brewing step, avoids possible water interference with the GC analysis, and results in a quicker analytical screening because of automation and shorter sampling procedure.

4. CONCLUSIONS

Coffee samples were analyzed with three sampling approaches (HS-SPME of the coffee powder, HS-SPME of the brew, and *in*-solution SBSE of the brew) coupled with GC-MS; each sampling can be considered as a part of the sensory experience perceived during cupping coffee evaluation. Despite the differences between the three sampling approaches, data processing showed that the three methods provide the same kind of chemical information useful for sample discrimination, and that they could be used interchangeably to sample the coffee aroma and Flavour. Comparison of the multivariate analysis of the sensory data with the chemical fingerprint of the investigated samples showed that: a) sensory and chemical data are in good agreement, and b) sensory evaluation can be related to the different chemical composition of the samples investigated. The choice of sampling technique used for this purpose may thus be guided by factors such as simplicity, sensitivity, reliability, and possibility of automation. As a consequence, HS-SPME of the coffee powder is the approach providing the most satisfactory performance, because: a) the direct sampling of coffee powder does not require further operations, while the brewing process may be a source of variability, b) HS-SPME affords full and easier automation of the analytical procedure, and c) HS-SPME of the coffee powder provides the largest number of identified (or tentatively identified) components.

Further in-depth studies will be necessary to correlate groups of compounds to a specific sensory note characterizing coffee samples, and to enable the development of a predictive model to support sensory panels in their sensory evaluation of coffee samples. In addition, knowledge on the odor active compounds correlated to a characteristic note, the concentrations of these compounds and their interactions, may open new perspectives in understanding the biological mechanisms underlying the pleasure related to the aroma and Flavour of coffee.

Acknowledgements
This study was supported by Ricerca Finanziata da Università - Fondo per la Ricerca Locale (Ex 60%) Anno2014.

REFERENCES

1. International Trade Centre. *The Coffee Exporter ' s Guide*. (2011).
2. Sunarharum, W. B., Williams, D. J. & Smyth, H. E. Complexity of coffee flavor: A compositional and sensory perspective. *Food Res. Int.* **62,** 315–325 (2014).
3. Blank, I., Sen, A. & Grosch, W. Potent odorants of the roasted powder and brew of Arabica coffee. *Leb. Und-forsch.* **195,** 239–245 (1992).
4. Czerny, M. & Grosch, W. Potent odorants of raw Arabica coffee. Their changes during roasting. *J. Agric. Food Chem.* **48,** 868–872 (2000).
5. Flament, I. *Coffee Flavor Chemistry*. (2001).
6. Frank, O., Zehentbauer, G. & Hofmann, T. Bioresponse-guided decomposition of roast coffee beverage and identification of key bitter taste compounds. *Eur. Food Res. Technol.* **222,** 492–508 (2006).
7. Picó, Y. *Chemical Analysis of Food: Techniques and Applications*. (Elsevier 2012)
8. Chiralertpong, A., Acree, T. E., Barnard, J. & Siebert, K. J. Taste–Odor Integration in Espresso Coffee. *Chemosens. Percept.* **1,** 147–152 (2008).
9. Chiralertpong, A., Acree, T. E., Barnard, J. & Siebert, K. J. *Taste-Odor Integration in Espresso Coffee*. *Chemosensory Perception* **1,** 147–152 (Springer New York LLC, 2008)
10. Köster, E. P., Köster, E. P. & Mojet, J. Theories of food choice development. *Underst. Consum. food Prod.* 93–124 (2007).
11. Prescott, J. Multisensory processes in flavour perception and their influence on food choice. *Curr. Opin. Food Sci.* **3,** 47–52 (2015).
12. Dunkel, A. *et al.* Nature's chemical signatures in human olfaction: A foodborne perspective for future biotechnology. *Angew. Chemie - Int. Ed.* **53,** 7124–7143 (2014).
13. Geithe, C., Andersen, G., Malki, A. & Krautwurst, D. A Butter Aroma Recombinate Activates Human Class-I Odorant Receptors. *J. Agric. Food Chem.* **63,** 9410–9420 (2015).
14. Michishita, T. *et al.* Gas Chromatography/Olfactometry and Electronic Nose Analyses of Retronasal Aroma of Espresso and Correlation with Sensory Evaluation by an Artificial Neural Network. *J. Food Sci.* **75,** S477–S489 (2010).
15. Bhumiratana, N., Adhikari, K. & Chambers, E. Evolution of sensory aroma attributes from coffee beans to brewed coffee. *LWT - Food Sci. Technol.* **44,** 2185–2192 (2011).
16. Ruosi, M. R. *et al.* A further tool to monitor the coffee roasting process: Aroma composition and chemical indices. *J. Agric. Food Chem.* **60,** 11283–11291 (2012).
17. Liberto, E. *et al.* Non-separative Headspace Solid Phase Microextraction–Mass Spectrometry Profile as a Marker To Monitor Coffee Roasting Degree. *J. Agric. Food Chem.* **61,** 1652–1660 (2013).
18. Ribeiro, J. S., Augusto, F., Salva, T. J. G. & Ferreira, M. M. C. Prediction models for Arabica coffee beverage quality based on aroma analyses and chemometrics. *Talanta* **101,** 253–260 (2012).
19. Ribeiro, J. S., Augusto, F., Salva, T. J. G., Thomaziello, R. A. & Ferreira, M. M. C. Prediction of sensory properties of Brazilian Arabica roasted coffees by headspace solid phase microextraction-gas chromatography and partial least squares. *Anal. Chim. Acta* **634,** 172–179 (2009).
20. Science, F., Pérez-Martínez, M., Sopelana, P., de Peña, M. P. & Cid, C. Application of Multivariate Analysis to the Effects of Stored Coffee Brew. *J. Agric. Food Chem.* **56,** 11845–11853 (2008).
21. Ongo, E. *et al.* Chemometric discrimination of philippine civet coffee using electronic nose and gas chromatography mass spectrometry. *Procedia Eng.* **47,** 977–980 (2012).
22. Lindinger, C. *et al.* When Machine Tastes Coffee: Instrumental Approach To Predict the

Sensory Profile of Espresso Coffee. *Anal. Chem.* **80,** 1574–1581 (2008).
23. Wang, Y., O'Reilly, J., Chen, Y. & Pawliszyn, J. Equilibrium in-fibre standardisation technique for solid-phase microextraction. *J. Chromatogr. A* **1072,** 13–17 (2005).
24. Bicchi, C., Cordero, C., Liberto, E., Sgorbini, B. & Rubiolo, P. Reliability of fibres in solid-phase microextraction for routine analysis of the headspace of aromatic and medicinal plants. *J. Chromatogr. A* **1152,** 138–149 (2007).
25. Sgorbini, B. *et al.* New medium-to-high polarity twister coatings for liquid and vapour phase sorptive extraction of matrices of vegetable origin. *J. Chromatogr. A* **1265,** 39–45 (2012).
26. Mestdagh, F., Davidek, T., Chaumonteuil, M., Folmer, B. & Blank, I. The kinetics of coffee aroma extraction. *Food Res. Int.* **63,** 271–274 (2014).

Further Reading

http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_tc_browse.htm?commid=47950S [Retreived October 2015]

http://www.scaa.org/PDF/resources/cupping-protocols.pdf [December 2015]

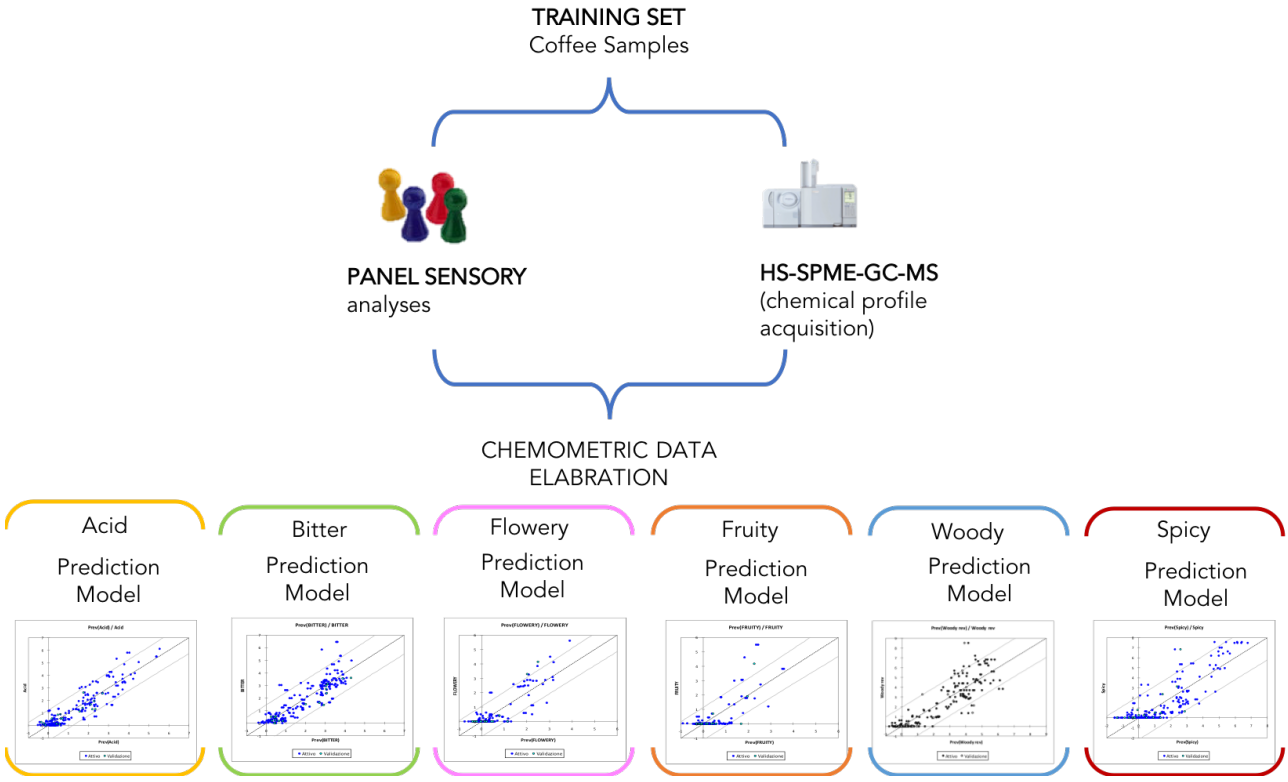
EPI Suite v3.10 developed by the EPA's Office of Pollution Prevention Toxics NS Syracuse Research Corporation (SRS) 2000 U.S.

2.2 Coffee sensory notes sensometric definition and inter-approach validation

This work has been presented as an oral communication (code YS14) at the Fourth International Congress on Cocoa Coffee and Tea (CoCoTea) 2017, Turin, Italy 25-28 June 2017
with the title:
"Coffee aroma sensory notes and chemical signature: a sensometric approach"

D. Bressanello*[1], E. Liberto[1], C. Cordero[1], B. Sgorbini[1], P. Rubiolo[1], G. Pellegrino[2], R. Ruosi[2], C. Bicchi[1]
[1]University of Turin, Italy, [2]Lavazza spa, Italy

0. GRAPHICAL ABSTRACT



1. INTRODUCTION

Coffee aroma and Flavour are the most important properties determining consumers' preference and acceptance, and, together with price, are the main properties underlying a coffee's commercial success [1–3]. Coffee companies need to evaluate the quality of the beans they are going to buy to determine the best coffees to use in their blends, and/or which coffees are of sufficiently high quality to be commercialized as "single origin". Several attributes are considered in qualifying a good coffee such as lack of defects, beans color and size and flavour but the latter evaluation is recognized as indispensable. Cupping protocols are international standards for cupping and grading coffees as a function of their sensory properties [4,5].

However, sensory methods are expensive and time-consuming in terms of panel training and alignment, and sometimes not sufficiently precise for a critical and objective evaluation. Last but not least it is rather difficult to apply at-line for immediate feedback. The sensory approach adopted so far has been quantitative descriptive analysis, with studies at the molecular level to disclose relationships between chemical composition and sensory response. The sensory lexicon related to coffee is a descriptive tool used worldwide to define aroma and Flavour attributes quantitatively, through scaled scores. The lexicon used when measuring the sensory aspects of different Flavour attributes is therefore of great importance and must carefully be defined when chemical and sensory data have to be compared, because non-specific language may create confusion. This approach has successfully been used by trained professionals to evaluate coffee, enabling different panels to obtain the same intensity score for each attribute for a given sample [6,7,8]. However, well-designed and standardized chemo-sensory evaluation is the key point to identify the chemicals responsible for a given sensorial note. The definition of a relationship between chemical profile and aroma sensory impact is thus an important challenge, in both the analytical and industrial fields, because it may enable food industries to obtain an objective evaluation (independent or complementary to the panel subjectivity) of their products.

Gas-chromatography combined with mass-spectrometry (GC-MS) is the analytical technique of election to study the composition of coffee aroma [9–11]Conversely, several sampling approaches have been used to extract and concentrate the Flavour components, and more in general volatile compounds, directly from the ground coffee (powder) and/or from the coffee brew. Sample preparation is the crucial step in any analytical procedure, and must enable to recover the chemical components representative of each sensory informative note [3,10,12–23] Furthermore, in quality control the whole analytical procedure used to study chemicals related to the sensory experience should be integrated and fully automated.

In this context, sensometrics acts as a bridge linking sensory properties to the chemical information underlying them; it can be used in quality assurance and control, in product/blend development, in benchmarking new products and to evaluate their probable market impact and in predicting preferences based on formulation changes [24,25].
Furthermore, the sensometric approach can provide information about the chemicals that discriminate the sensory attributes and link them to sensory perception through correlation models. The conventional strategy in aroma studies implies that a single odorant is considered together with its sensory description. On the opposite, the correlation through sensometrics enables: i) to measure all compounds eliciting a peculiar sensory perception simultaneously, and ii) to link the quali-quantitative distribution of odorants and their mutual and cross-modal interactions to the odor perceived, through the sensory scores provided by the panel [13,18,22,26,27]. However, reliable models that describe a sensory note representative of the variability of coffee

require: i) a large number of different coffee samples, and ii) fast analytical techniques also applicable at-line or on-line, as a complement to the decision of the sensory panel.
Different studies have addressed the sensory-instrumental relationship as regards the sensory properties of coffee, but many of these looked at correlations between small and homogeneous pool of samples, without giving suitable representativeness to the wide variability of coffees due to origin, post-harvest processing, and roasting [18,28,29]).
This study is part of a wider project aiming at correlating the sensory characteristics of coffee aroma with its chemical composition, in order to provide an objective tool complementary to sensory evaluation, and to be applied for routine use.

This part of the study, in particular, focuses on the use of sensometrics as a tool to define the odorants characterizing the chemical signature of different coffee aroma notes and to validate it for sensory scores prediction (Figure 1). Coffee samples with peculiar sensory characteristics of different origins, species, and submitted to different post-harvest treatments, have been included as representative of commercial coffees as it is in quality control at the industrial level in the acceptance in incoming beans; specimens were analysed both sensorially and for their volatiles composition. Sensory evaluation was done by an expert coffee-cupping panel, through a quantitative descriptive analysis using a monadic approach. Sensory attributes included acidity, bitterness, woody, fruity, flowery, spicy, and nutty notes, aroma intensity, body, and astringency. Chemical analyses were carried out by Headspace Solid-Phase Micro-Extraction combined with gas chromatography mass spectrometry (HS-SPME-GC-MS).
This technique may also be included within an automatic Total Analysis System (TAS), with which a large number of samples can be screened for quality control of in-cup coffee sensory quality[9,30,31].

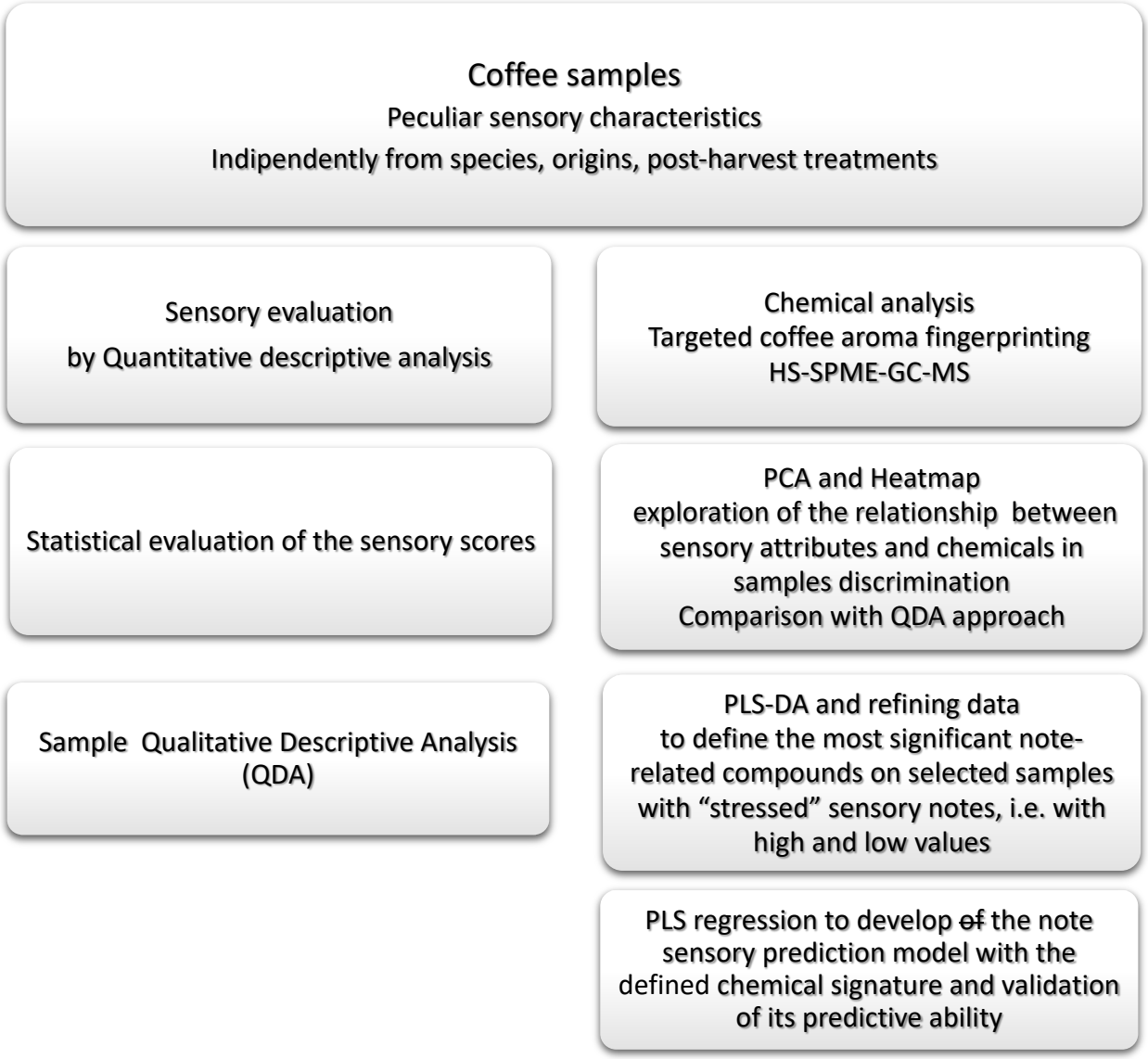


Figure 1 Sensometric data treatment workflow in revealing the signature of the coffee aroma notes

2. MATERIALS AND METHODS

2.1 Reagents and Matrices.

Lavazza Spa (Turin, Italy) kindly supplied coffee samples, consisting of roasted ground coffee suitable for a coffee-filter machine, over a period of 15 months.

157 coffee samples with distinctive sensory notes, originating from different countries (Ethiopia, Papua New Guinea, Colombia, Brazil, India, Indonesia, Tanzania, Uganda, and Vietnam), of the species *Coffea arabica L.* (Arabica) and *Coffea canephora Pierre* (Robusta), were analysed in replicates. Table 1APX (in the appendix at page 19) reports the complete list of samples investigated in this study. Samples 1-8 were used during the first part of the study to determine which sample preparation method was the most suitable to study the relationship between chemical and sensory attributes (Bressanello et al. 2017). Samples 9 to 20 were a selection of samples specifically characterized by woody and bitter notes, with scores at the two extremes (highest-lowest) of a 0-10 scale. The roasting degree of each sample was carefully measured by ground bean light reflectance, with a single-beam Neuhaus Neotec Color Test II instrument (Genderkesee, Germany) at a wavelength of 900 nm, on 25-30g of ground coffee. Roasting degree was set at 55°Nh, to be close to the international standardization protocol for cupping[5] samples were roasted within 24 hours prior to cupping, and left for at least 8 hours to stabilize. For clarity, samples are labelled in the text with their origins. Pure reference standards for identity confirmation acetic acid, 3-methyl butanoic acid, 3-methyl butanal, 2-methyl butanal, 2,3-Dimethylpyrazine, 2,3,5-Trimethylpyrazine, 2,5-Dimethyl pyrazine, 2-Ethyl-3-methylpyrazine, 2,5-Dimethyl-3(2H)-furanone, 2-ethyl-3,6-dimethyl pyrazine, 2–methyl pyrazined, 2-Furanmethanethiol, 5-Methyl furfural, Furfural, furfuryl acohol, Fufuryl disulfide, Guaiacol, Ethyl guaiacol, Vinyl guaiacol, Maltol and n-alkanes (n-C9 to n-C25) for Linear Retention Index (ITS) determination were from Sigma-Aldrich (Milan, Italy).

Internal standards (ISTDs) for analyte response normalization were n-C13. A standard stock solution of ISTDs at 1000 mg/L was prepared in dibuthylphtalate (Sigma-Aldrich, Milan, Italy) and stored in a sealed vial at -18°C.

2.2 Headspace solid phase microextraction (HS-SPME) sampling

The fiber chosen was a Polydimethylsiloxane/Divinylbenzene (PDMS/DVB) df 65 µm, 1 cm length; it was conditioned before use as recommended by the manufacturer. The SPME device was from Supelco (Bellfonte, PA, USA). Coffee aroma compounds were sampled by automated Headspace Solid Phase Microextraction (auto-HS-SPME) using a Combi-PAL AOC 5000 (Shimadzu, Milan, Italy) assembled on-line in a Shimadzu QP2010 GC–MS system provided with Shimadzu GC–MS Solution 2.51 software (Shimadzu, Milan, Italy).

HS-SPME of the coffee powder: 1.500 ± 0.010 g of powder were weighed in a septum-sealed gas vial (20 mL); the resulting headspace was sampled through the PDMS/DVB SPME fiber for 40 minutes at 50°C at a stirring speed of 350rpm. The internal standard had previously been loaded onto the fiber[32] by sampling 5L of a 1000 mg/L solution of n-C13 in DBP in a 20mL headspace vial for 20 min at 50°C, stirring speed 350rpm. After sampling, the accumulated analytes were recovered by thermal desorption of the fiber for 5 min at 250°C into the GC injector, and then transferred on-line to the gas-chromatographic column. All samples were analysed in duplicate.

2.3 GC-MS analysis conditions

GC-MS analysis - Chromatographic conditions: injector temperature: 250°C, injection mode: splitless; carrier gas: helium, flow rate: 1 mL/min; fiber desorption and reconditioning times: 5 min; column: SGE SolGelwax (100% polyethylene glycol) 30 m x 0.25 mm dc x 0.25 µm df (SGE-Melbourne, Australia). Temperature program, from 40°C (1min) to 200°C at 3°C/min, then to 250°C (5min) at 10°C/min.

MSD conditions: ionization mode: EI (70 eV), temperatures: ion source: 200°C; quadrupole: 150°C; transfer line: 250°C; scan range: 35-350 m/z.

2.4 Identification of volatile components.

Aroma compounds sampled from the headspace of the coffee powder were identified through their linear retention indices (ITS) and EI-MS spectra, compared to those of authentic standards or, tentatively, to those collected in-house or in commercial libraries (Wiley 7N and Nist 05 ver 2.0 Mass Spectral Data).

2.6 Descriptive sensory analysis of coffee aroma

The samples were submitted to sensory evaluation through a quantitative descriptive analysis by the Lavazza trained panel expert in coffee evaluation following the Q cupping and grading protocol according to SCAA [4,5,33]). The protocol entails three tasting steps, after roasting to a set colour (55-60° Nh) and eight hours of sample stabilization: i) evaluation of the aroma by sniffing the dry ground coffee, ii) evaluation of the aroma by sniffing the brew three minutes after its preparation and stirring, and iii) Flavour evaluation after 8-10 minutes. The attributes aftertaste, acidity, body, and balance are evaluated by tasting the brew, spraying it into the mouth to maximize retro-nasal vapours. Cup quality was assessed for several attributes; those considered in this study were: Flavour (flowery, fruity, woody, nutty, spicy), acidity, bitterness, body (mouthfeel), astringency, and aroma intensity. The quality and intensity of each attribute were evaluated simultaneously, upon a scale from 0 to 10. ANOVA analysis was run to verify the panel alignment on each attribute under consideration. Averages from experts who had similarity were used as “main scores” for attributes under investigation.

2.7 Data processing

Chromatographic data were collected using Shimadzu GCMS Solution 2.5SU1 software.
Principal Component Analysis (PCA) was used to detect sample groups and outlier(s) within chemical and sensory data. Partial Least Square Discriminant Analysis (PLS-DA) was then performed on the sensory scaled samples (low-high score range) to identify the compounds most closely related to a sensory attribute, and Partial Least Square Regression (PLS) was used to correlate chemicals to sensory attributes, and to evaluate the ability of extracted chemical variables to predict sensory scores. HS-SPME-GC-MS profiles normalized to ISTD were used, and auto-scaling was applied as data pretreatment: this step ensured that the contribution of each X variable (odorants) to the Y variable (sensory scores) was unbiased. One-way ANOVA and t-test on the sensorial results, and PCA, PLS-DA, and PLS, were run by XLSTAT (version 2015.5.01.23164) software, copyright Addinsoft 1995-2015.

3. RESULTS AND DISCUSSION

3.1 Qualitative Descriptive Analysis (QDA)

PCA is relevant in sensometrics mainly as a standard tool for Qualitative Descriptive Analysis (QDA). In QDA, a panel of trained assessors rates a number of samples for perceived intensities of distinct attributes on scales, according to reference protocols for specific food commodities (e.g. coffee, olive oil), depending on the panel's experience and/or on the complexity of the matrix. By averaging these intensity ratings and replicates, it is possible to build up a data matrix, in which the rows are food samples, and the columns the relative sensory attributes used to describe them [24,25]. Analysis of this data matrix by PCA can give information both on how coffee samples are related, and on which sensory notes best describe each sample. PCA was applied to the mean QDA sensory scores for aroma and Flavour of eight samples analyzed in five replicates by five assessors; the bi-plot of scores and loadings is shown in **Figure 2**. The PCs that accounted for 75.4% (PC1) and 14.4% (PC2) of the total variance were extracted. Focusing on sensory attributes, aroma intensity dominated Robusta samples (JAV, UGA INDO), and appeared to be correlated to spicy, woody, body, and bitter notes. Acid and bitter are normal taste attributes; however, previous studies demonstrated that correlations between volatiles and taste sensory attributes can be found, since several volatiles and non-volatiles have common reaction pathways during roasting [18,28,27].
Moreover, panelists perceive odors via ortho- and retro-nasal pathways, as a result of both compound mutual interactions and of cross-modal effects between odorants and taste that can amplify or modify perception, which, physiologically, occurs in the brain. Conversely, this interaction does not occur at the molecular level [34–36].
Bitter notes are also closely related to nutty and astringent notes. In contrast, fruity, flowery and acid showed vectors different from those of the above descriptors, and were positively correlated with Arabica samples (COL, PNG, INDIA, KAFA, BRA). Among these, the only exception is the India sample (INDIA ARAB CHERRY) that shows sensory characteristics more similar to Robusta samples.

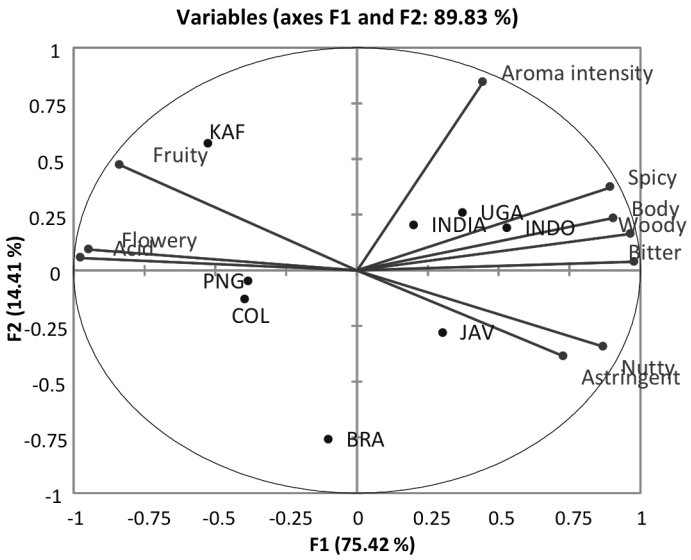


Figure 2 Biplot of scores and factor loadings obtained by PCA for aroma descriptors of QDA for coffees 1-8 (n=40).

3.2 From sensory evaluation to the related chemicals: a discriminative and informative guide

A total of 95 compounds were identified (or tentatively identified; 17 compounds were unknown (or not unequivocally identified) by HS-SPME-GC-MS in the coffee powder samples. Table 1 reports the list of identified compounds with their Linear Retention Indices (I^Ts). The coffee aroma chemical profile of the first 40 samples obtained by HS-SPME-GC-MS was processed by PCA, together with their sensory scores, to determine: a) whether groups and/or outliers were present, and b) the relationship between samples and chemical-sensory variables. As expected, as well as QDA analysis, PCA on the aroma chemical profile showed a discrimination ability, firstly driven by species (Figure 3A), and then within species, by the sensory characteristics peculiar of each origin (Figure 3C-E). The Loadings Plot (Figure 3B) clearly shows that the sensory notes are split into two groups: 1) the acid, flowery, fruity notes, which are located in the 1st quadrant of the Cartesian plane, i.e. the same as the Arabica samples, 2) the bitter, nutty, woody, and spicy notes, lying in the 3rd quadrant, i.e. the same as the Robusta samples. Several chemical variables are linked to this sample distribution, and thus to these sensory notes, but they may be present in more than one note (Figure 3B**)**.

Several pyrazines, (e.g 2-*n*-propylpyrazine, 2,6-diethylpyrazine, 2-methyl-3,5-diethylpyrazine, isopropenylpyrazine) and phenolic derivatives, such as guaiacols, characterize the Robusta samples and are more closely related to the roasty, tobacco, nutty, spicy, and woody notes, while furan derivatives, esters and ketones are linked to the sweet, fruity, and floral sensory attributes [9,20,37].

Within the Robusta samples, a) JAVA is the most nutty sample, characterized by the compounds #48 (unknowns 4) b) UGA samples have high acidity and are chemically described by 2,3-butandione and 2,3-pentandione, acetoxyacetone, hexanal, acetic acid, 1-hydroxy-2-butanone, and 1H-pyrrole-2-carboxaldehyde, which elicit musty, sour, pungent, buttery notes that can be related to acid attributes, c) INDO is the most woody, spicy, and bitter sample, and is more full-bodied and astringent, mostly characterized by several pyrazines and phenolic compounds, as previously mentioned.

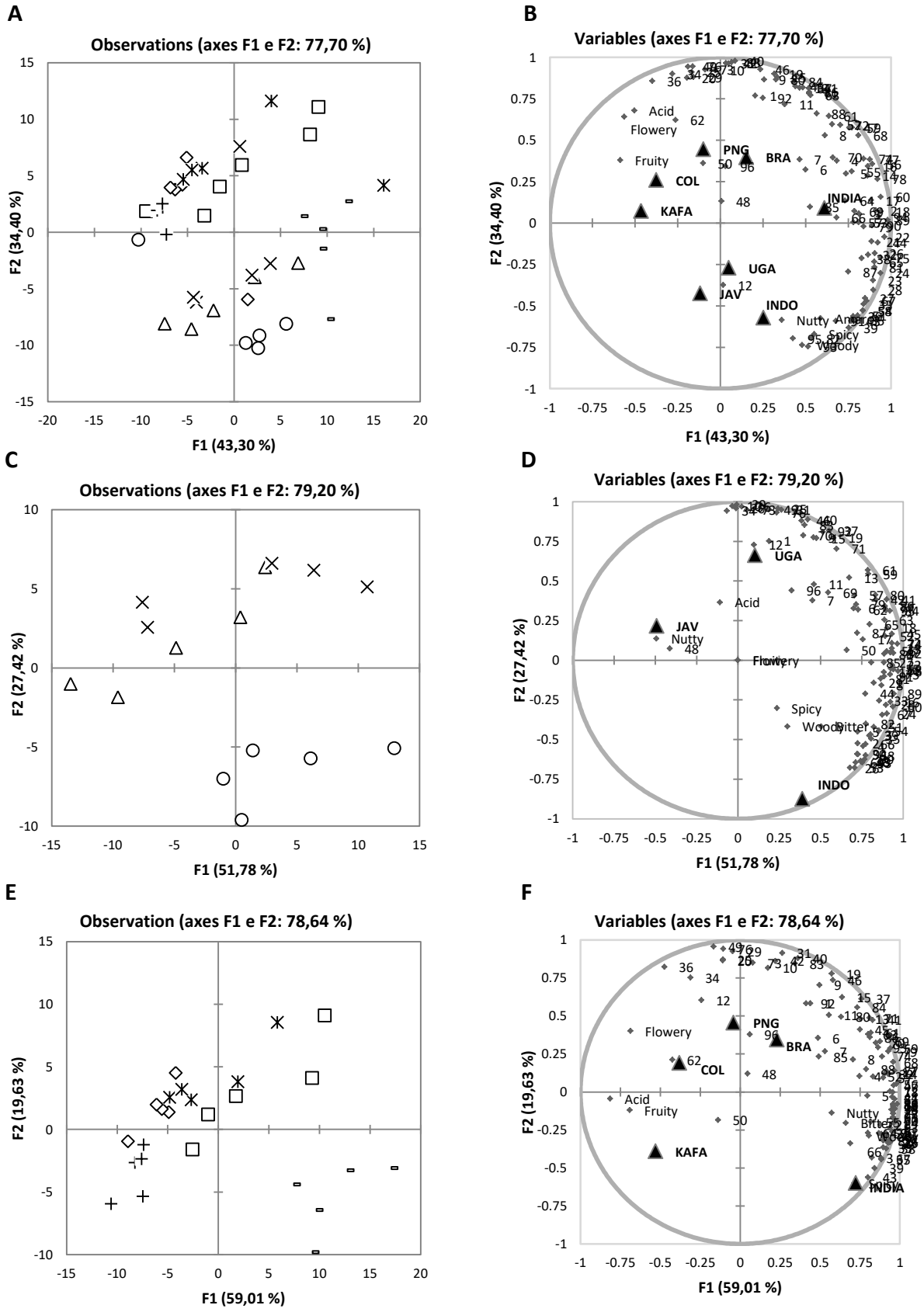


Figure 3 Scores and loading plots of the coffee samples obtained by chemical and sensory analysis, A&B all samples, C&D Robusta ; E&F Arabica coffees. Legend: BRA: □; COL: ◊; JAV: Δ; UGA: X; PNG: * ; INDIA: ▭ ; INDO: ◯; KAFA: +

Within Arabica, separation between samples is achieved on PC1 and PC2: a) KAFAs are characterized by high body, fruity note, and aroma intensity; these sensory attributes are chemically described by furfuryl alcohol, methyl acetate, 5-methyl furfural, 2-cyclopenten-1-one-3 methyl, all of which are characterized by sweet, fruity, malty, and nutty notes, b) conversely, PNG is mostly characterized by acetyl furan, 2-furfuryl-5-methylfurane, 2-furanmethanol propanoate, 2-furfuryl furan, eliciting more fruity and floral attributes. Although HS-SPME implies discrimination of the analytes depending on their volatility/polarity, if used under standardized conditions it provides reliable information for fingerprinting studies and perfectly suitable for comparative analyses [38]. However, despite the great informative potential of PCA applied to comprehensive sensory and chemical data, for the purpose of discriminating samples by their sensory characteristics, it is still challenging to define the chemical fingerprint of a note. It is therefore necessary to analyze the relationship between chemical compounds and sensory note descriptions in greater depth, looking at the chemical variables most closely correlated with each sensory note. Thus, the Pearson correlation coefficient was applied to the PCA correlation matrix for each sensory note, and only variables with a value above 0.5 were selected. Table 1 reports the variables related to each sensory attribute, (except for aroma intensity, body, and astringency), for which the Pearson coefficient was above the cut-off (taken at $r \geq 0.5$). Table 1 shows that specific sensory notes (e.g. acid and flowery, or woody, bitter, nutty and spicy) are often described by the same variables, i.e. the components closely correlated with these notes are very often the same. These results confirm that a specific sensory note is described by component amounts and ratios, and rarely by single specific compounds [17]. This may be due to the complexity of odor and taste perception and to their mutual influence on the actual perceived Flavour. Interactions among odorants gives odor synesthesia while interaction between odorants and tastants could give chemesthetic sensations[1,2]. Moreover, several examples have been reported showing that some compounds have odors directly related to the note (X-L) and others not (X), i.e., when alone they elicit a different sensorial perception [34]. The percentage contribution of each compound to the whole-aroma chemical profile closely correlated to a given sensory attribute was monitored across all samples investigated, to study the relationship of the percentage contribution to the expression of a given sensory characteristic.

Table 1 List of identified and *tentatively identified compounds. (http://webbook.nist.gov/chemistry/name-ser.html). Compounds with an "X" are related to each sensory note; "L" indicates that the compound alone directly elicits the peculiar note perception. Odour description is deriving from literature data (°http://www.thegoodscentscompany.com/, & www.flavornet.org, + Blank et al.).

	Compound Name	Odour Description°,&,+	Calc. I^Ts	Lit. I^Ts	Acid	Fruity	Flowery	Bitter	Nutty	Woody	Spicy
1	Acetaldehyde	Fruity	706	723							
2	Acetone	Ethereal, Apple, Pear	824	835							
3	Methyl acetate	Ethereal, Sweet, Fruity	828	839							
4	2-methylfuran	Chocolate, Nutty	885	864							
5	2-Butanone	Ethereal	906	905							
6	2-Methylbutanal	Chocolate	914	931							
7	3-Methylbutanal	Aldehydic	918	936							
8	2,5-dimethylfuran	Meaty	950	939							
9	2,3-Butanedione	Buttery	978	963							
10	2,3-Pentanedione	Buttery	1058	1060	X		X				
11	2-Vinylfuran	Nutty, coffee	1071	1085							
12	Hexanal	Green	1080	1098							
13	2,3-Hexanedione	Buttery	1130	1110							
14	1-methyl-1H-Pyrrole	Woody	1137	1140							
15	2-Vinyl-5-methylfuran	-	1151	1152							
16	Pyridine	Fishy	1177	1177							
17	Pyrazine	Sweet, Floral	1209	1206				X			
18	Methylpyrazine	Nutty	1262	1268				X			
19	3-Hydroxy-2-butanone	Buttery	1281	1285							
20	1-Hydroxy-2-propanone	Sweet-Caramellic	1297	1318	X		X				
21	2,5-Dimethylpyrazine	Nutty-Roasted	1317	1321				X		X	X
22	2,6-Dimethylpyrazine	Earthy-Chocolaty	1324	1327				X		X	X
23	Ethylpyrazine	Nutty-roasted	1329	1343				X	X-L	X	X
24	2,3-Dimethylpyrazine	Green, Nutty	1341	1354				X	X-L	X	X
25	1-Hydroxy-2-butanone	Sweet-caramellic	1370	1381	X		X				

	Compound Name	Odour Description°,&,+	Calc. I^Ts	Lit. I^Ts	Acid	Fruity	Flowery	Bitter	Nutty	Woody	Spicy
26	3-Ethylpyridine	Tobacco	1374	1384				X		X	X
27	2-Ethyl-5-methylpyrazine	Coffee-like	1386	1399				X	X-L	X	X
28	2-Ethyl-3-methylpyrazine+ Trimethylpyrazine	Nutty	1399	1400				X	X-L	X	X
29	UNK 1(m/z: 54 [100%]; 43[78%]; 42[16.75%])		1405	-	X		X				
30	2-n-Propylpyrazine *	Nutty	1413	1425				X	X-L	X	X
31	UNK 2(m/z: 112 [100%]; 68[73.76%]; 40[24.93%])		1418	-	X		X				
32	2-Furanmethanethiol	Roasted, Burnt	1432	1440				X	X	X-L	X
33	2-Ethyl-3,6-dimethylpyrazine	Earthy, Baked	1441	1449				X-L	X	X	X
34	Acetic acid	Sour, Pungent	1446	1454	X-L	X-L					
35	2,6-Diethylpyrazine	Nutty	1457	1458				X	X-L	X	X
36	Furfural	Sweet, Woody, Bready	1462	1467	X		X				
37	Acetoxyacetone	Fruity	1468	1467							
38	Furfurylmethylsulfide	Sulphuraceous, Garlic	1483	1496				X		X	X
39	3,5-Diethyl-2-methylpyrazine	Nutty	1489	1491				X	X	X-L	X
40	2,5-Dimethyl-3(2H)-furanone	Caramellic	1496	1490	X		X-L				
41	Furfuryl formate	Ethereal	1497	1497							
42	Acetylfuran	Sweet-caramellic	1500	1498	X		X				
43	3-Methyl-2-Cyclopenten-1-one + 3,5-Diethyl-2-methylpyrazine	Nutty	1509	1509				X	X	X-L	X
44	1-H-Pyrrole	Nutty	1513	1525				X	X	X-L	X
45	Belzaldehyde + 2-Methyl-3(2H)-thiophenone *+ Furan-2-yl-propan-2-ol		1519	1520							
46	2-Oxopropylpropanoate	-	1531	1531							
47	Furfuryl acetate	Sweet, Fruity, Banana	1538	1539							
48	UNK 4(m/z: 110[100%]; 109[86.72%]; 53[50.99%])		1564	-							
49	5-Methyl Furfural	Caramellic	1570	1562	X		X				

	Compound Name	Odour Description°,&,+	Calc. I^Ts	Lit. I^Ts	Acid	Fruity	Flowery	Bitter	Nutty	Woody	Spicy
50	2,3-Butandiole + 1-(5-Methyl-2-furyl)2-propanone		1575	1582							
51	1-methylethenylpyrazine	Roasted, Nutty	1590	-				X	X-L	X	X
52	Furfuryl propanoate	Fruity	1598	1603							
53	2-Furfurylfuran	-	1608	-							X
54	(5H)-5-Methyl-6,7-dihydrocyclopentapyrazine	Earthy	1611	1611				X	X-L	X	X
55	1-Methylpyrrole-2-carboxaldehyde	Nutty	1614	1635							
56	4-Hydroxybutanoate	-	1621	-							
57	UNK 6 (m/z: 137[100%]; 94[61.57%]; 122[37.71%])		1630	-							
58	2-Isopropenylpyrazine	Caramellic, Nutty	1633	1633				X	X-L	X	X
59	2,5-Dihydro-3,5-dimethyl-2-furanone *	-	1642	1640							
60	1-(1-methyl-1H-pyrrol-2-yl) – ethanone *+2-Acetyl-5-methyl pyrrole	-	1649	-							
61	Furfurylalcohol	Bready	1661	1664							
62	3-Methylbutanoic acid	Acid, Fruity, Sour	1667	1670	X-L	X-L	X-L				
63	3-Methyl-1,2-ciclohexanedione	-	1678	-							
64	2-Furfuryl-5-methylfuran	-	1686	1636							
65	2-Acetyl-3-methylpyrazine*	Nutty	1694	1719				X	X-L	X	X
66	Furfurylpentanoate		1702	1702							
67	2-Methyl-6-(1-propenyl)-pyrazine*	Chocolate, Nutty	1708	1719				X	X-L	X	X
68	UNK 11 (m/z:69[84-54%]; 41[100%];83 m/z[31.67%])		1709	-							
69	1-Acetyl-1,4-dihydropyridine*		1716	-				X			
70	UNK 12 (m/z: 140[100%]; 43[55%]; 111[33.52%])		1726	-							
71	UNK 13 (m/z: 54[100%]; 82[73.67%] ; 110 [57.91%])		1729	-							

	Compound Name	Odour Description°,&,+	Calc. I^Ts	Lit. I^Ts	Acid	Fruity	Flowery	Bitter	Nutty	Woody	Spicy
72	UNK 13b (m/z : 67 [100%]; 112[73.67%]; 53[55.59%])		1734	-							
73	UNK 14 (m/z: 55; 84[48.75%]; 54[26.45%])		1745	-	X		X				
74	UNK 15 (m/z: 119[100%]; 43[26.78%]; 64[25.39%])		1750	-							
75	Methyl nicotinate + other		1767	1778				X		X	X
76	UNK 17 (m/z: 95[100%]; 43[28.81%]; 138 [17.56%])		1772	-	X		X				
77	UNK 18 (m/z: 123; 122[74.34%]; 126 [11.84%])		1777	-							
78	3-Methyl-2-butenoic acid	Phenolic	1786	-							
79	UNK 20 (m/z:139; 43[12.25%]; 154[50.48%])		1791	-				X			
80	2-Hydroxy-3-methyl-2-cyclopenten-1-one	Malty	1820	1839							
81	Furfurylpyrrole	Vegetable	1853	1866				X		X	X
82	Guaiacol	Spicy	1860	-				X		X	X-L
83	2-Acetyl-5-methylfuran	Nutty	1885	-	X		X				
84	3-Ethyl-2-hydroxy-2-Cyclopenten-1-one	Caramellic	1885	-							
85	trans-Furfurylideneacetone*	-	1897	-							
86	Phenylethanol	Floral	1902	1912				X		X	X
87	Maltol	Sweet, Caramell	1952	1960							
88	2-Acetylpyrrole	Musty	1962	1971				X			X
89	Difurfurylether	Nutty, Earthy	1980	1977				X			X
90	Phenol deriv*	-	1997	-				X		X	X
91	1H-Pyrrole-2-carboxaldehyde	Musty	2012	2035							
92	4-Ethyl-guaiacol	Spicy	2021	2037				X		X	X-L
93	Nonanoic acid	Waxy, Fatty	2150	2159							
94	4-Vinyl-guaiacol	Woody	2185	2193				X		X-L	X
95	Difurfuryldisulfide	Coffee	2536	2536							

Figure 4 shows a “heat map” of the samples, scored on the normalized percentage contributions of components closely correlated with woody, nutty, and “fresh” notes (acid, flowery/fruity). The slots in each row are colored according to the magnitude of their values, from red (low percentage) to yellow (high percentage); for instance, guaiacol mainly contributes to the profiles of INDO, UGA and JAVA samples. INDO samples had the highest contribution from variables related to woody. INDIA samples, despite being Arabica, show sensory characteristics similar to Robusta, thus confirming the sensory scores given by the panel from the chemical standpoint (Table 1APX). Conversely, guaiacol, 1H-pyrrole-2-carboxaldehyde, (5H)-5-methyl-6,7-dihydrocyclopentan pyrazine, 2-furanmethanethiol, and difurfuryldisulfide are directly involved in defining the woody note [17,39]

A

Woody	BRA	COL	JAV	UGA	PNG	INDIA	INDO	KAFA
2,5-Dimethylpyrazine	2.87	2.67	3.64	3.4	2.54	3.52	3.98	3.45
2,6-Dimethylpyrazine	3.46	3.09	4.51	4.02	3.13	4.3	4.68	3.53
Ethylpyrazine	1.44	1.21	2.22	1.96	1.22	1.78	2.19	1.35
2,3-Dimethylpyrazine	0.64	0.51	0.9	0.82	0.54	0.89	1.08	0.55
3-Ethylpyridine	0.12	0.08	0.12	0.13	0.1	0.2	0.23	0.1
2-Ethyl-5-methylpyrazine	1.26	1.12	2.23	1.89	1.09	1.78	2.26	1.37
2-Ethyl-3-methylpyrazine+ Trimethylpyrazine	1.59	1.35	2.64	2.27	1.36	2.46	2.98	1.57
2-n-Propylpyrazine	0.05	0.04	0.09	0.07	0.04	0.07	0.11	0.04
2-Furanmethanethiol	0.29	0.22	0.4	0.38	0.24	0.31	0.43	0.18
2-Ethyl-3,6-dimethylpyrazine	0.91	0.75	1.84	1.5	0.66	1.65	1.99	0.66
2,6-Diethylpyrazine	0.35	0.29	0.73	0.6	0.3	0.62	0.9	0.36
methyl nicotinate + others	0.1	0.08	0.13	0.12	0.08	0.12	0.14	0.07
Furfurylmethylsulphide	0.17	0.19	0.2	0.24	0.22	0.3	0.42	0.2
3,5-Diethyl-2-methylpyrazine	0.15	0.12	0.39	0.31	0.13	0.32	0.45	0.17
3-Methyl-2-Cyclopenten-1-one + 3,5-Diethyl-2-methylpyrazine	0.04	0.03	0.09	0.08	0.03	0.1	0.14	0.04
1-H-Pyrrole	0.54	0.34	0.64	0.57	0.4	0.65	0.69	0.33
1-methylethenylpyrazine	0.15	0.12	0.28	0.22	0.12	0.21	0.32	0.13
(5H)-5-Methyl-6,7-dihydrocyclopentapyrazine	0.2	0.17	0.37	0.33	0.18	0.32	0.46	0.19
1-Methylpyrrole-2-carboxaldehyde	0.88	0.74	0.82	0.89	0.86	0.94	0.98	1.09
2-Isopropenylpyrazine	0.23	0.17	0.42	0.37	0.19	0.41	0.57	0.2
2-Acetyl-3-methylpyrazine	0.37	0.31	0.56	0.46	0.31	0.45	0.5	0.34
2-Methyl-6-(1-propenyl)-pyrazine	0.18	0.15	0.35	0.29	0.15	0.34	0.4	0.18
Furfurylpyrrole	1.05	1.05	1.67	1.37	1.12	1.32	1.71	1.03
Guaiacol	0.38	0.42	1.01	1.56	0.48	0.74	2.02	0.44
Maltol	0.16	0.1	0.14	0.32	0.13	0.19	0.27	0.09
1H-Pyrrole-2-carboxaldehyde	0.27	0.29	0.53	0.66	0.3	0.4	0.71	0.29
Nonanoic acid	0.32	0.3	1.19	1.25	0.3	0.61	2.22	0.17
Difurfuryldisulfide	1.83	1.94	5.54	5.35	1.61	1.88	5.4	1.02

B

Nutty	BRA	COL	JAV	UGA	PNG	INDIA	INDO	KAFA
Ethylpyrazine	1.44	1.21	2.22	1.96	1.22	1.78	2.19	1.35
2,3-Dimethylpyrazine	0.64	0.51	0.9	0.82	0.54	0.89	1.08	0.55
2-Ethyl-5-methylpyrazine	1.26	1.12	2.23	1.89	1.09	1.78	2.26	1.37
2-Ethyl-3-methylpyrazine+ Trimethylpyrazine	1.59	1.35	2.64	2.27	1.36	2.46	2.98	1.57
2-n-Propylpyrazine	0.05	0.04	0.09	0.07	0.04	0.07	0.11	0.04
2-Ethyl-3,6-dimethylpyrazine	0.91	0.75	1.84	1.5	0.66	1.65	1.99	0.66
2,6-Diethylpyrazine	0.35	0.29	0.73	0.6	0.3	0.62	0.9	0.36
3,5-Diethyl-2-methylpyrazine	0.15	0.12	0.39	0.31	0.13	0.32	0.45	0.17
3-Methyl-2-Cyclopenten-1-one + 3,5-Diethyl-2-methylpyrazine	0.04	0.03	0.09	0.08	0.03	0.1	0.14	0.04
1-H-Pyrrole	0.54	0.34	0.64	0.57	0.4	0.65	0.69	0.33
1-methylethenylpyrazine	0.15	0.12	0.28	0.22	0.12	0.21	0.32	0.13
Furfuryl propanoate	0.29	0.22	0.4	0.38	0.24	0.31	0.43	0.18
(5H)-5-Methyl-6,7-dihydrocyclopentapyrazine	0.2	0.17	0.37	0.33	0.18	0.32	0.46	0.19
2-Isopropenylpyrazine	0.23	0.17	0.42	0.37	0.19	0.41	0.57	0.2
2-Acetyl-3-methylpyrazine	0.37	0.31	0.56	0.46	0.31	0.45	0.5	0.34
2-Methyl-6-(1-propenyl)-pyrazine	0.18	0.15	0.35	0.29	0.15	0.34	0.4	0.18
Methyl nicotinate + other	0.1	0.08	0.13	0.12	0.08	0.12	0.14	0.07
Guaiacol	0.38	0.42	1.01	1.56	0.48	0.74	2.02	0.44
Maltol	0.16	0.1	0.14	0.32	0.13	0.19	0.27	0.09
1H-Pyrrole-2-carboxaldehyde	0.27	0.29	0.53	0.66	0.3	0.4	0.71	0.29
Nonanoic acid	0.32	0.3	1.19	1.25	0.3	0.61	2.22	0.17
Difurfuryldisulfide	1.83	1.94	5.54	5.35	1.61	1.88	5.4	1.02

C

Acid/Fuity/Flowery	BRA	COL	JAV	UGA	PNG	INDIA	INDO	KAFA
2,3-Pentanedione	0.3	0.34	0.19	0.18	0.35	0.19	0.09	0.31
1-Hydroxy-2-propanone	0.86	1.12	0.58	0.54	0.92	0.48	0.3	0.77
1-Hydroxy-2-butanone	0.08	0.1	0.04	0.05	0.08	0.04	0.03	0.07
UNK 1(m/z: 54 [100%]; 43[78%]; 42[16.75%])	0.38	0.45	0.25	0.24	0.37	0.18	0.11	0.33
UNK 2(m/z: 112 [100%]; 68[73.76%]; 40[24.93%])	0.13	0.14	0.09	0.1	0.14	0.09	0.06	0.14
Acetic acid	4.78	7.64	3.54	3.16	7.25	3.83	2.04	6.85
Furfural	3.24	5.37	2	2.33	4.13	1.6	1.02	4.02
2,5-Dimethyl-3(2H)-furanone	0.64	0.73	0.48	0.52	0.74	0.5	0.36	0.74
Acetylfuran	1.39	1.89	0.93	1.1	1.81	1.2	0.99	1.78
5-Methyl Furfural	6.02	7.8	4.48	5.19	7.15	3.99	2.88	7.35
3-Methylbutanoic acid	2.35	3.07	2.14	2.71	3.53	2.03	2.41	5.15
UNK 14 (m/z: 55; 84[48.75%]; 54[26.45%])	0.26	0.33	0.19	0.17	0.3	0.19	0.12	0.27
UNK 17 (m/z: 95[100%]; 43[28.81%]; 138 [17.56%])	0.34	0.47	0.26	0.29	0.43	0.24	0.18	0.42
2-Acetyl-5-methylfuran	0.83	1.03	0.56	0.67	1.07	0.7	0.41	0.94

Figure **4** Percentage contribution of compounds correlated with: A) Woody, B) Nutty and C) Fresh notes (Acid, Flowery, Fruity)

Little may be said about the compounds linked to the fruity, acid, and flowery notes, because of the similarity of the chemicals involved. However, these compounds give a higher contribution to Arabica than to Robusta samples [16,19,28]), and COL contained their highest concentrations within the Arabica group. The variables found to be correlated to fruity, acid, and flowery notes are in full agreement with the sensory scores of those samples (Table 1APX). For instance, fruity related compounds, such as acetic acid and 3-methylbutanoic acid, are massively present in KAFA samples, which achieved the highest score for the fruity note. (Figure 4 c)

Two aspects at present make it difficult to understand which compound specifically contributes to a given note, and how it contributes to it. The first is that the chemical definition of a sensory note (i.e. its aroma signature) of a sample is linked to its composition, not only qualitatively but also quantitatively, and, in particular, to the ratios between components. The second aspect is reflected in the narrow range of the scores of some notes, e.g. for nutty, from 0 to 3. When the range is very limited, seeking odorants to be correlated to the sensory note becomes challenging. To overcome these limitations, it is therefore necessary to maximize the range of sensory scores, by selecting those samples having "stressed" sensory notes, i.e. with high and low values, so as to define more precisely the aroma compounds describing the note, and to verify the method ability to correlate them with the sensory fingerprint.

A selection of "stressed"[2] sensory samples representative of each note considered, independently of species, origin, and post-harvest treatment, were analyzed and the chemical findings related to the sensory scores (Table 1APX).

[2] "Stressed" samples means a panel selection of a new pool of samples with considerable differences of sensory score within a given note

3.2.1 Definition of note-related compounds, on representative "stressed" sensory samples

A supervised chemometric tool (PLS-DA) was applied to study note-related compounds. PLS-DA describes samples by calculating new variables that maximize separation between groups while minimizing variability within groups. Samples with the lowest score (for each target note) were assigned to class 1, while those with the highest score were assigned to class 2. The impact of each compound on the separation of the pool of samples into the two classes (1 and 2) was evaluated by VIP (Variable Impact on Projections). The cut-off was arbitrarily chosen, for each note, as the point at which the VIP values dropped sharply in the histogram.

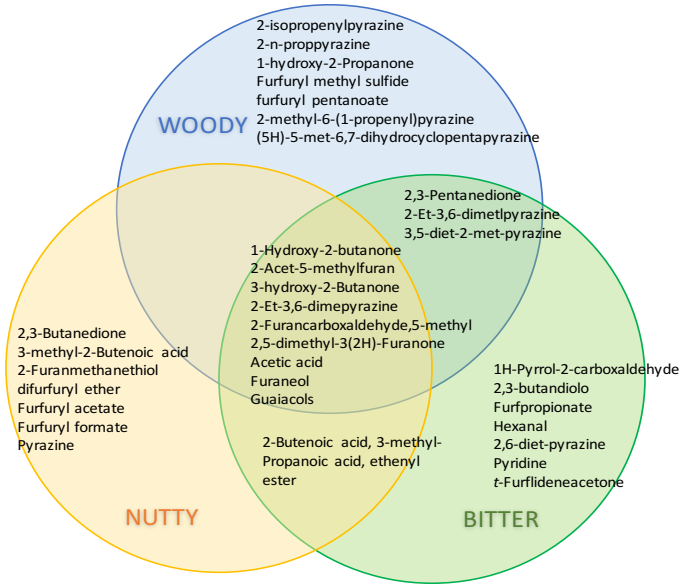


Figure 5 Venn Diagram of note-related compounds extracted from PLS-DA and the refining step and used to build the note prediction model for woody, bitter, and nutty attributes.

Figure 5 reports the compounds related to woody, bitter, and nutty notes, identified by the above procedure, and then used to build the prediction models. Some points were raised by this selection:

a) a single compound can contribute to the score of more than one sensory note, as reported by Ribeiro et al. in a study on prediction models of the quality of Arabica coffee beverages[27];

b) the variables selected to describe the woody note are probably a consequence of clearer definition (in terms of maximized scores and in panel alignment on the lexicon) of the note across the group of samples, which led to more precise identification of the most significant note-related compounds.

3.3 From chemistry to senses: sensometric driver tool for sensory notes prediction

The compounds identified with PLS-DA for each sensory attribute (Figure 5) were used to study the correlation between chemical and sensory data on the whole data set, through a note-prediction model with Partial Least Square Regression (PLS). Samples were randomly divided into three groups: a training set (131), a validation set (10), and a test set (15). Figure 6 lists the parameters used to build the model (A), together with the regression curve and validation set fit (B) and the ability to predict the sensory scores for the woody note on test samples (C). The results show correct sample distribution across the calibration interval, showing that the model is reliable for sensory score prediction. Prediction reliability was evaluated through deviation (Residues) from the predicted *vs.* experimental scores, calculated as follows [1]:

$$Res. = Experimental\ Score - Predicted\ Score \qquad [1]$$

The model showed close correlation between odorants selected and sensory scores. The predictive ability was good, i.e. Q^2= 0.66, with a prediction error of 0.59. Similar results, although to a different extent, were obtained for the other sensory attributes (Figure 1APX). The only exception was the nutty note, for which the model showed poor linearity (R^2=0.467), a wide confidence limit (RMSE= 2.709) and very low predictive ability (mean error: 1.595), mainly due to the difficulty in the lexicon used to define nutty and, as a consequence, to determine odorants linked to it.

Although to different extents, the results show good relationships between selected odorants and sensory scores, with an average standard error in score prediction of ± 1, thus cross-validating the link between the compounds selected and the sensory note description.

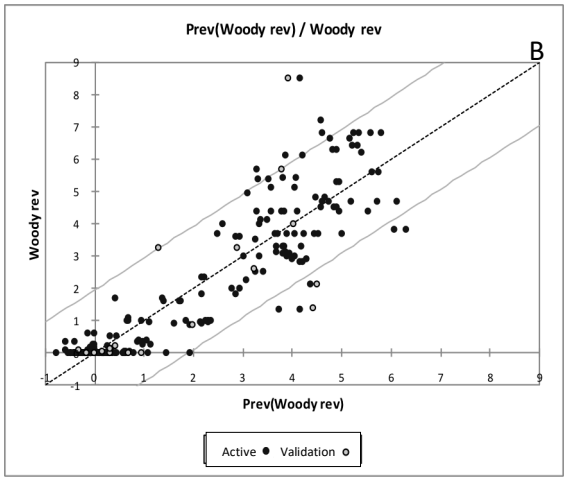
In this perspective, the chemical composition of woody and flowery notes were also investigated by sensomics [40,41]. The preliminary results of this inter-approach validation showed good consistency between sensometrics and sensomics for some of the target compounds. The compounds already confirmed by sensomics are listed in Figure 7. The sensometric-driven procedure can reveal the chemical aroma signature of the sensory attributes investigated. At present, both performance and predictive ability of the models are too closely related to the training set, and to its ability to cover the entire range of scores of the samples under study, i.e. woody *vs.* fruity. Compared to fruity, the woody note offers better prediction, thanks to its robustness, owing to two factors: the wide pool of samples, covering the whole sensory score range and thus avoiding the need to force the model; and the good panel alignment in the woody note lexicon recognition, providing a better estimate of an external test set. [26–28,42]

Although this discussion has focused on woody, nutty, and bitter notes, acid, spicy, and flowery/fruity notes were also considered (figure 1APX); they acted similarly to the notes discussed in depth, including positive aspects and limitations.

A

Training Set		
Data Matrix	n° Observations	218
	n° of Explicative Variables	23
	n° of Quantitative Dependent Variables	1
Cross Validation	n° Random Chosen Observations	18
Test Set	n° Random Chosen Observations	30

B



C

Observation	Predicted	Measured	Res. (Abs Value)
1_1	-0,42	0,00	0,42
12_1	-0,35	0,00	0,35
13_1	0,10	0,00	0,10
35_2	0,07	0,00	0,07
5_1	0,34	0,00	0,34
50_2	-0,14	0,00	0,14
6_1	0,53	0,00	0,53
7_1	0,25	0,00	0,25
74_1	-0,77	0,00	0,77
82_2	-0,35	0,00	0,35
85_2	-0,21	0,00	0,21
25_2	0,30	0,03	0,27
45_1	0,80	0,03	0,77
61_1	-0,04	0,03	0,06
1_BRALA2_AN_2	0,73	1,00	0,27
53_1	0,67	1,07	0,40
89_2	3,75	1,40	2,35
1_JAV_RL_1	1,74	1,60	0,14
87_1	2,69	2,25	0,44
3_KAAP_RL_2	3,27	2,60	0,67
79_2	3,91	3,53	0,38
39_1	3,37	3,70	0,33
1_CON_RN_2	4,34	4,00	0,34
54_2	4,10	4,40	0,30
3_BUK_RN_1	4,65	4,50	0,15
20_1	4,40	4,68	0,28
57_2	3,08	4,95	1,87
65_1	5,02	6,20	1,18
1_INDOEK1_RT_1	4,12	6,80	2,68
1_UGA_RN_2	4,67	7,20	2,53

Model Performance Parameters	
R^2	0.816
Q^2	0.659
Median of Error in Prediction	0.343

Figure **6** A) Parameters used to build the model, B) regression curve and validation set fit, C) results of sensory score prediction on the test samples for the woody note.

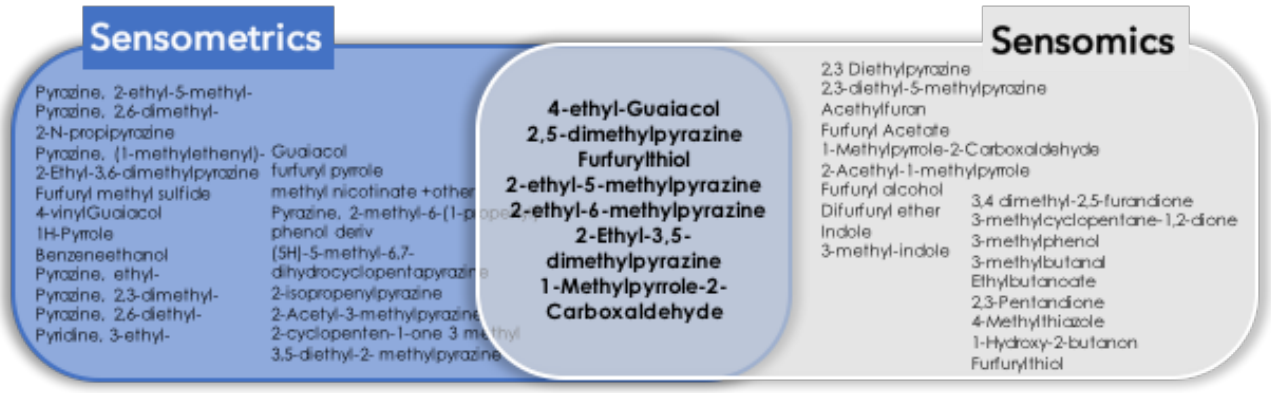


Figure **7** Compounds characterizing woody and flowery aroma notes confirmed by sensomics.

4. CONCLUSIONS

The sensometric-driven approach has been found to be discriminative, informative, and predictive in revealing the chemical signature of the different coffee aroma notes: a) discriminative, because it was able to point out samples with peculiar aroma notes , independently of species, post-harvest treatment, and origins, b) informative, because it entails the complementary and simultaneous use of sensory and chemical data to define odorants able to describe the chemistry of aroma notes, stressing the method strengths and limitations, and c) predictive, because the panel-coherent sensory score prediction based on this sensometric approach confirms and reinforces the relevance and significance of the odorants selected by applying this procedure. The models developed are promising for predicting sensory scores from chemical data and appear to provide a complementary tool that can contribute to an objective sensory evaluation, despite the great variability of coffee samples (origins, species, treatments, qualities), that aim to reflect what happens at the industrial level in the selection of the raw beans to be used as monorigin or in blends.

These results are a good background to build up a reliable mathematical model for a future application to routine quality control although further experiments are necessary on a more extended number of samples to improve the representativeness of the treated coffees, and a panel's alignment on a more specific lexicon to define the notes.

Acknowledgements

This study was supported by Ricerca Finanziata da Università - Fondo per la Ricerca Locale (Ex 60%) Anno2014.

Davide Bressanello is grateful to Prof. Dr. Peter Schieberle of the DFA (Deutsche Forschungsanstalt für Lebensmittelchemie) for hospitality and the time spent at the DFA, and for fruitful discussion on the sensomic approach.

REFERENCES:

1. Chiralertpong, A., Acree, T. E., Barnard, J. & Siebert, K. J. Taste–Odor Integration in Espresso Coffee. *Chemosens. Percept.* **1,** 147–152 (2008).
2. Prescott, J. Multisensory processes in flavour perception and their influence on food choice. *Curr. Opin. Food Sci.* **3,** 47–52 (2015).
3. Sunarharum, W. B., Williams, D. J. & Smyth, H. E. Complexity of coffee flavor: A compositional and sensory perspective. *Food Res. Int.* **62,** 315–325 (2014).
4. Britta Folmer. The Craft and Science of Coffee. in *The Craft and Science of Coffee* (eds. Blank, I., Farah, A., Giuliano, P., Sanders, D. & Wille, C.) (Elsevier, 2017).
5. SCAA. *SCAA Protocols Cupping Specialty Coffee*. *Specialty Coffee Association of America* (2015).
6. WCR. WCR (World Coffee Research). Available at: https://worldcoffeeresearch.org/. (Accessed: 26th August 2016)
7. Kubota, L. Reinventing the (Flavor) Wheel: Industry Collaborates to Identify Coffee Flavor. *The Specialty Coffee Chronicle*
8. Chambers, E. *et al.* Development of a 'living' lexicon for descriptive sensory analysis of brewed coffee. *J. Sens. Stud.* **31,** 465–480 (2016).
9. Liberto, E. *et al.* Non-separative Headspace Solid Phase Microextraction–Mass Spectrometry Profile as a Marker To Monitor Coffee Roasting Degree. *J. Agric. Food Chem.* **61,** 1652–1660 (2013).
10. Pico, Y. *Chemical Analysis of Food Techniques ans Application*. *Elsevier* (2012).
11. Buettner, A. *Handbook of Odor*. (Springer, 2017).
12. Bhumiratana, N., Adhikari, K. & Chambers, E. Evolution of sensory aroma attributes from coffee beans to brewed coffee. *LWT - Food Sci. Technol.* **44,** 2185–2192 (2011).
13. Blank, I., Alina, S. & Grosch, W. Potent odorants of the roasted powder and brew of Arabica coffee. *Zeitschrift fur Leb. Und-forsch.* **195,** 239–245 (1992).
14. Bressanello, D. *et al.* Coffee aroma : Chemometric comparison of the chemical information provided by three different samplings combined with GC – MS to describe the sensory properties in cup. *Food Chem.* **214,** 218–226 (2017).
15. Czerny, M. & Grosch, W. Potent odorants of raw Arabica coffee. Their changes during roasting. *J. Agric. Food Chem.* **48,** 868–872 (2000).
16. Flament, I. *Coffee Flavor Chemistry*. (Wiley 2001).
17. Folmer, B. How can science help to create new value in coffee? *Food Res. Int.* **63,** 477–482 (2014).
18. Lindinger, C. *et al.* When Machine Tastes Coffee: Instrumental Approach To Predict the Sensory Profile of Espresso Coffee. *Anal. Chem.* **80,** 1574–1581 (2008).
19. Mayer, F. & Grosch, W. Aroma simulation on the basis of the odourant composition of roasted coffee headspace. *Flavour Fragr. J.* **16,** 180–190 (2001).
20. Ruosi, M. R. *et al.* A Further Tool To Monitor the Coffee Roasting Process: Aroma Composition and Chemical Indices. *J. Agric. Food Chem.* **60,** 11283–11291 (2012).
21. Scheidig, C., Czerny, M. & Schieberle, P. Changes in key odorants of raw coffee beans during storage under defined conditions. *J. Agric. Food Chem.* **55,** 5768–5775 (2007).
22. Semmelroch, P. & Grosch, W. Studies on Character Impact Odorants of Coffee Brews. *J. Agric. Food Chem.* **44,** 537–543 (1996).
23. Semmelroch, P. & Grosch, W. Analysis of roasted coffee powders and brews by gas chromatography-olfactometry of headspace samples. *LWT - Food Sci. Technol.* **28,** 310–313 (1995).

24. Ochi, H. *et al.* Sensometric calibration of sensory characteristics of commercially available milk products with instrumental data. *J. Dairy Sci.* **93,** 1794–806 (2010).
25. Tomic, O. Using sensometrics to extract information from your sensory data. *Food Sci. Technol.* **26,** 42–44 (2012).
26. Pongsuwan, W. *et al.* Prediction of Japanese green tea ranking by gas chromatography/mass spectrometry-based hydrophilic metabolite fingerprinting. *J. Agric. Food Chem.* **55,** 231–236 (2007).
27. Ribeiro, J. S., Augusto, F., Salva, T. J. G. & Ferreira, M. M. C. Prediction models for Arabica coffee beverage quality based on aroma analyses and chemometrics. *Talanta* **101,** 253–260 (2012).
28. Ribeiro, J. S., Augusto, F., Salva, T. J. G., Thomaziello, R. A. & Ferreira, M. M. C. Prediction of sensory properties of Brazilian Arabica roasted coffees by headspace solid phase microextraction-gas chromatography and partial least squares. *Anal. Chim. Acta* **634,** 172–179 (2009).
29. Ribeiro, J. S., Ferreira, M. M. C. & Salva, T. J. G. Chemometric models for the quantitative descriptive sensory analysis of Arabica coffee beverages using near infrared spectroscopy. *Talanta* **83,** 1352–1358 (2011).
30. Cagliero, C. *et al.* Fast headspace-enantioselective GC–mass spectrometric-multivariate statistical method for routine authentication of flavoured fruit foods. *Food Chem.* **132,** 1071–1079 (2012).
31. Manz, A., Graber, N. & Widmer, H. M. Miniaturized total chemical analysis systems: A novel concept for chemical sensing. *Sensors Actuators B Chem.* **1,** 244–248 (1990).
32. Wang, Y., O'Reilly, J., Chen, Y. & Pawliszyn, J. Equilibrium in-fibre standardisation technique for solid-phase microextraction. *J. Chromatogr. A* **1072,** 13–17 (2005).
33. Lawless, H. T. & Heymann, H. *Sensory Evaluation of Food*. *Science* (Springer New York, 2010).
34. Chambers IV, E. & Koppel, K. Associations of volatile compounds with sensory aroma and flavor: The complex nature of flavor. *Molecules* **18,** 4887–4905 (2013).
35. Donishi, Kimura, A., Imbe, H., Yokoi, I. & Kaneoke, Y. Sub-threshold cross-modal sensory interaction in the thalamus: Lemniscal auditory response in the medial geniculate nucleus is modulated by somatosensory stimulation. *Neuroscience* **174,** 200–215 (2011).
36. Dalton, P., Doolittle, N., Nagata, H. & Breslin, P. a. The merging of the senses: Integration of subthreshold taste and smell. *Nat. Neurosci.* **3,** 431–432 (2000).
37. Grosch, W. Flavour of coffee. A review. *Nahrung-Food* **42,** 344–350 (1998).
38. Bicchi, C., Cordero, C., Liberto, E., Sgorbini, B. & Rubiolo, P. Headspace Sampling in Flavor and Fragrance Field. in *Comprehensive Sampling and Sample Preparation* **4,** 1–25 (Elsevier, 2012).
39. Frank, O., Zehentbauer, G. & Hofmann, T. Bioresponse-guided decomposition of roast coffee beverage and identification of key bitter taste compounds. *Eur. Food Res. Technol.* **222,** 492–508 (2006).
40. Schieberle, P. New Developments in Methods for Analysis of Volatile Flavor Compounds and their Precursors. in *Characterization of Food: Emerging Methods* 403–431 (1995).
41. Schieberle, P. & Hofmann, T. Mapping the combinatorial code of food flavors by means of molecular sensory science approach. in *Food flavors; chemical, sensory, and technological properties.* (ed. Henryk, J.) 413–438 (CRC Press 2011)
42. Ikeda, T., Kanaya, S., Yonetani, T., Kobayashi, A. & Fukusaki, E. Prediction of Japanese Green Tea Ranking by Fourier Transform Near-Infrared Reflectance Spectroscopy. *J. Agric. Food Chem.* **55,** 9908–9912 (2007).

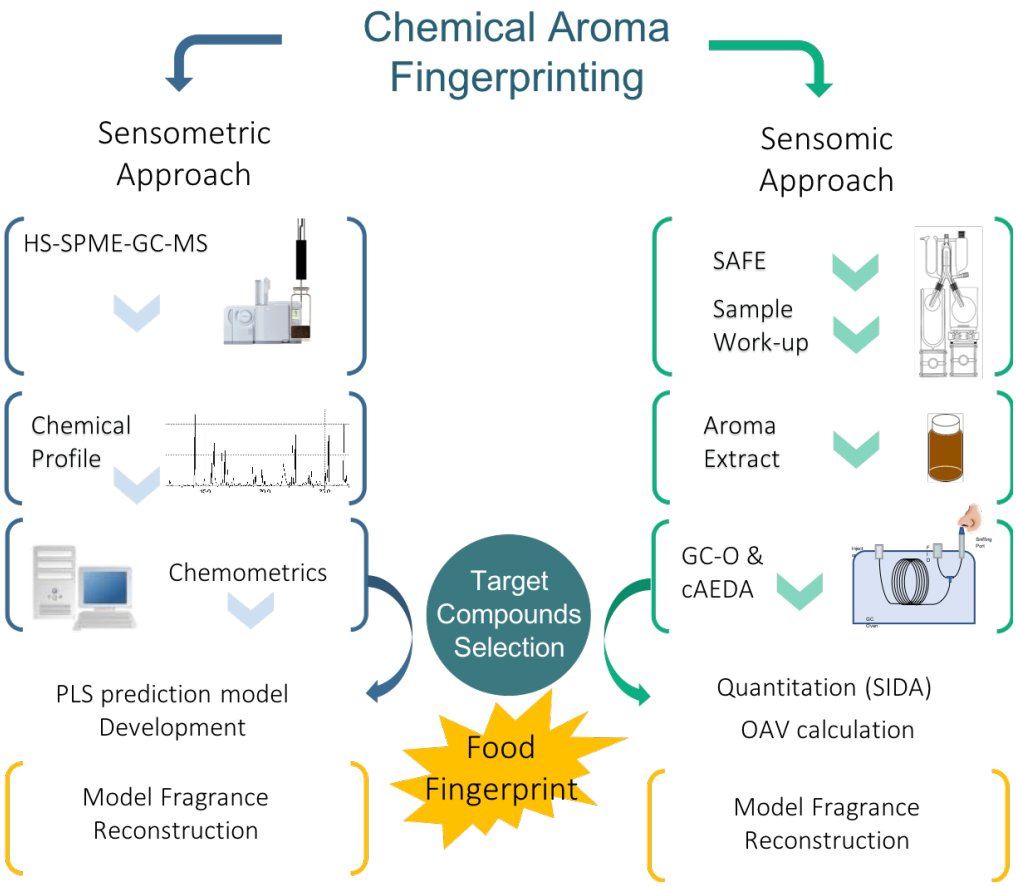
2.3 Sensometric and molecular sensory science: inter-approach validation

This work has been presented as a poster (code P1B.14) at the Fourth International Congress on Cocoa Coffee and Tea (CoCoTea) 2017, Turin, Italy 25-28 June 2017
with the title:
"Comparison between sensometrics and sensomic approaches in the sensory-chemistry relationship definition"

Davide Bressanello[1], Erica Liberto[1]*, Chiara Cordero[1], Barbara Sgorbini[1], Cecilia Cagliero[1] Patrizia Rubiolo[1], Gloria Pellegrino[2], Rosanna M. Ruosi[1], Carlo Bicchi[1]

1 Dipartimento di Scienza e Tecnologia del Farmaco, Università degli Studi di Torino, Via Pietro Giuria 9, Torino, Italy
2Lavazza spa, strada Settimo 410, 10156 Torino, Italy

0. GRAPHICAL ABSTRACT



1. INTRODUCTION AND OBJECTIVES

The sensometric approach, developed in the previous steps of the project seemed promising in the definition of a correlation between sensory data and the chemical profiles of coffee samples. The fast and automatable TAS affords the screening of a high number of samples (mandatory in quality control) while the use of statistical tools (e.g. PLS-DA and PLS) makes possible to connect the sensory evaluation of the panel with the chemical profile obtained with the HS-SPME-GC-MS platform.
Nevertheless, for a reliable validation, the results of this approach must be compared to those of molecular sensory science approach i.e. the reference approach adopted to identify and quantify the molecules responsible for different foods Flavours.

The aim of this study is to apply the sensomic approach to selected coffee samples able to minimize and maximize both "*Woody*" and "*Flowery*" sensory notes. Key aroma compounds identified with the molecular sensory science approach were compared to those selected by sensometrics called Note Related Compounds (NRC) and used to develop the prediction models.

This part of the project has been carried out at the Deutsche Forschungsanstalt für Lebensmittelchemie (DFA) of the Technische Universität München (TUM) under the supervision of Prof. Dr. Peter Schieberle

Strategy for Molecular Sensory Science (or Sensomics)

The molecular sensory science is also called "sensomics" since, as other –omic approaches, it includes a biological data dimension represented by the sensory perception and it is largely adopted in food aroma investigations.
At the basis of the molecular sensory science approach there is the statement: "not all volatile compounds are actually responsible for the food aroma". Developed by W. Grosh, P. Schieberle and their groups at the technical University of Munich, molecular sensory science became the gold standard for food aroma investigations[1–3,4,5,6].
The idea behind this approach is simple: extract aroma compounds from the matrix, separate and screen the mixture by GC-O/FID to select only the compounds with the highest odor impact and try to create with them a recombinant model able to elicit the same aroma of the food under study.
The molecular sensory science is constituted by five main steps summarized in figure 1

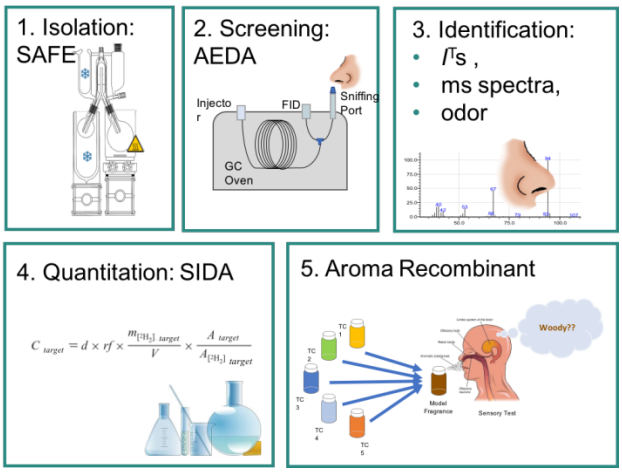


Figure 1 Summary of the steps of the Molecular Sensory Science approach

Isolation

Volatiles isolation is usually performed by Solvent Assisted Flavour Evaporation (SAFE). This technique, developed and introduced by Engel et al in 1999[7], is an evolution of simultaneous distillation-distillation (SDE) that overcomes its main limit linked to the high temperatures applied during the extraction that may lead to artefacts formation.
According to its forefathers the SAFE affords to:

- Keep the structures of the compounds of interest thanks to the low operative temperatures,
- Avoid interference from the non-volatile fraction,
- Provide a non-discriminant volatile extraction.

Concerning this last statement, since the aroma compounds are extracted by a solvent, a certain degree of discrimination has to be considered. Among the pool of volatiles, those with a higher affinity with the solvent will easily be extracted compared to low affinity components.
A schematic view of the SAFE equipment is reported in figure 2

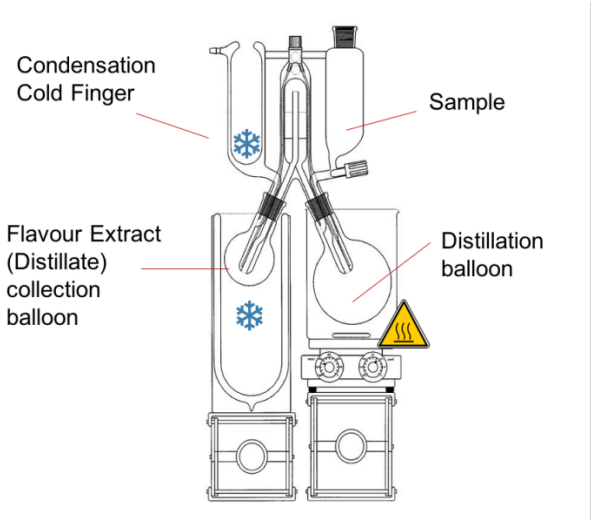


Figure 2 Scheme of the SAFE equipment.

Once the system is installed and perfectly sealed, the body and the legs of the distillation unit are termostated between 20 and 30°C with a warm water flow; this step avoids the condensation of the volatiles in undesirable parts of the distillation unit. The distillation balloon is then heated in a water bath between 20 and 30°C.
Then, the equipment is put under high vacuum (10^{-3}Pa) by a diffusion pump.
The condensation cold finger is hence filled by liquid nitrogen (-196°C) and the collection balloon immerged in liquid nitrogen.
Now the distillation can start; the sample is poured into the funnel and by dropping small aliquots of it into the distillation balloon from where the volatiles are recovered.
From the vapor spray, the volatiles are transferred to the distillation head where two glass propeller-shaped barriers are mounted to remove non-volatile compounds from the vapors. Then the volatile fraction reaches the cold trap, condenses and drops in the collection balloon.[7]

Screening: GC-O/FID and AEDA

This step is the core of the sensomic strategy since the key aroma candidates are identified by GC-O/FID within the Aroma Extract Dilution Analysis (AEDA) approach.
The SAFE distillate after being concentrated by micro-distillation is analyzed by Gas *Chromatography coupled to Olfactometry and Flame Ionization Detector (GC-O/FID).*
The analytical platform (described in section 1.3) implies a GC separation where, at the end of the column, the flow is splitted in two branches one directed to the sniffing port and the other to the FID detector that records the chromatogram.
During each GC run, the panellist places his nose closely above the top of the sniffing port and evaluates the odour of the chromatographic effluent. When an odour is perceived, the odour quality is assigned to the corresponding peak in the chromatogram. The accuracy of the peaks detection and the odours recognition increased with the experience, the panellist's training of the on pure reference compounds. This technique allows the combination of the sensory data of each odorant with its retention index (I^Ts).
Despite its good potentialities, GC-O has some important drawbacks; the number of compounds detected depends on many factors such as the panellist's experience and their concentration and perception threshold. Furthermore, in a single GC-O run is not possible to evaluate the intensity of an odour with a *single* sniffing and it is therefore not possible to understand if a compound is a key odorant or slightly contributes to the food aroma. These limitations are overcome by the Aroma Extract Dilution Analysis. This technique, developed by Grosch and Schieberle[2,1] affords to determine the relative odour activity of each compound without knowing its chemical structure.
AEDA is performed as follows: the volatile fraction is diluted stepwise 1:1 with solvent and each dilution analysed by GC-O/FID. The dilutions are continued until no odorants can be perceived in the GC effluent. The highest dilution to which each odorant can be perceived is called Flavour Dilution (FD) factor.
The results of the AEDA are usually displayed as a diagram of the FD-factors *vs.* the I^Ts called FD-chromatogram or Aromagram. This diagram makes possible to identify easily which peaks are more relevant in the definition of the overall aroma since they are characterized by a higher FD factor.
Comparative AEDA (cAEDA) was here applied because the objective of this study was to point out the differences between two different samples with the highest and the lowest expression of two different sensory notes the. The cAEDA found application also in the detection of off-flavour within different food matrices. In this procedure, peaks are not only selected according to their FD factor but also considering the difference of FD factors between the samples under study. Peaks with high FD factor differences represent the compounds responsible for the sensory differences between the samples.
The cAEDA has been performed as follow; the volatile fractions have been extracted from the same amount of both samples, worked-up, distilled and concentrated up to the same final volume. Assuming the same loss of volatiles in the samples preparation the comparison of single odorants in both samples on the basis of their FD-factors indicates which odorants are mainly responsible for the aroma differences. This comparison is done by comparing the aromagram in which each smelled peak is reported, and its retention index plotted *vs*. the FD factor assigned in each sample.
In agreement with the literature, only peaks with difference of at least 2 dilution steps are considered. This cut-off is necessary to prevent the operator variability.

Peaks Identification

The identification of the selected peaks is performed on the basis of three different criteria: I^Ts, Mass Spectrum and Odour Quality. The first step included the selection of some best candidates from a list of more than 1300 odour active compounds. Within this list, only compounds with a I^Ts and odour description compatible with those of the target peak on the aromagram are selected.
Then both the samples are injected on a GCxGC-TOF-MS system and the mass spectra are acquired. These analyses are useful to refine the first candidates list, to confirm some of them and find out other possible compounds responsible for target peaks. The identification is not only based on the similarity with commercial spectral libraries but also by comparison with Pure Referent Compounds.

The third step of the identification includes the investigation of the target odour quality and potency. The pure reference compounds are injected on the GC-O system and their odour quality compared to the odour description of target peaks on the AEDA. The identification of some critical compounds can also be checked by diluting the Pure Reference Compound up to a concentration similar to that one in which the FD factor was assigned in order to evaluate both the odour description and the potency.

Quantitation: Stable Isotopes Dilution Assay (SIDA) and determination of Odour Activity Values (OAV)

The importance of each odorant in the food aroma is evaluated by the calculation of their Odour Activity Values (OAV).
Before introducing the concept of OAV, it is important to define the Odour Threshold (OT): the Odour threshold is the minimum concentration at which the odour of a specific compound, in a given matrix (usually water, oil or air) can be perceived.
The importance of this parameter, reviewed by Ternanishi et al [8], indicates the odour potency of a specific compound; the smaller OT is the higher is the odour potency.
The OAV is defined as the ratio between the concentration of a compound in the matrix under study and its odour threshold as reported in the equation below:

$$OAV = \frac{[]_{target}}{OT_{target}}$$

Therefore, the absolute concentrations of the analytes have to be determined besides the odour thresholds. In molecular sensory science, the analytes concentration is determined by the Stable Isotopes Dilution Assay (SIDA) quantitation approach. The SIDA, introduced by Schieberle et al. [1], is a quantification method similar to the Standard Addition (SA) technique in witch isotopically labelled standards are used to determine analytes concentrations.
This method uses isotope-labelled derivatives of the investigated components (in general, ^{2}D or ^{13}C labelled compounds) as Internal Standards. They have the same physical-chemical properties of the investigated analytes and equal or very similar response factors and recoveries. Labelled compounds give also the same MS fragmentation pattern, but with a known increase in molecular weight, easily detectable by MS.[9]
the fragments selected for the quantitation may have different MS-responses, despite analytes and standards have the same properties. The MS-response factors (R_f) must therefore be calculated to compensate these differences.

A calibration line with mixtures of different analytes/standard ratios (1+5, 1+3, 1+1, 3+1, 5+1) is built up and analysed on the GC-MS system. The R_f is calculated using the following formula:

$$R_f = \frac{m_{ana} * A_{Std}}{m_{Std} * A_{Ana}}$$

whereas:
R_f= response factors.
m_{Ana}= amount of unlabelled analyte,
m_{Std}= amount of labelled standard,
A_{Ana}= peak area of ion m/z of unlabelled analyte,
A_{Std}= peak area of ion or sum of ions m/z of isotopically labelled standard

Samples are spiked with labelled compounds and then worked-up; with this procedure, all the losses in the sample preparation can be normalized in the quantitation.
The extracts are analysed on a GC-MS platform were peak areas of the labelled standard and of the analyte are determined only with the m/z fragments of interest.
The concentrations of the analytes in the samples are calculated with the following formula:

$$C_{Ana} = R_f \frac{m_{Std} * A_{Ana}}{g * A_{Std}}$$

whereas:
C_{Ana}= Concentration of the analyte,
R_f= response factors,
m_{Std}= amount of labelled standard,
g= amount of sample,
A_{Ana}= peak area of ion m/z of unlabelled analyte,
A_{Std}= peak area of ion or sum of ions m/z of isotopically labelled standard.

Aroma Recombinant

This step is mandatory for the full application of the molecular sensory science approach.
A model fragrance containing all the selected targets in proper concentration is built and its aroma is checked through sensory tests.
The model fragrance aroma has to be as close as possible to the aroma of the foodstuff under investigation; in order to refine this mixture, omission tests are also performed.
In omission tests, single or multiple analytes are removed from the mixture and the aroma is stepwise checked by sensory analysis to understand which are the so called "*key aroma compoun*ds", in other words those necessary to replicate the original aroma correctly.
In this session, the creation of the recombinant was a bit out of the main task but further studies should include the formulation of these kind of mixtures.

2. MATERIALS AND METHODS

2.1 Samples Description

Coffee samples has been selected according to their sensory profiles determined by the Lavazza's sensory panel. The samples were selected to maximize the expression of the sensory notes of interest (*Woody* and *Flowery*).

Coffees samples, consisting of roasted coffee ground to suit a coffee-filter machine were kindly supplied over a period of 9 months by Lavazza Srl (Turin, Italy).

The *Woody* note was from a Vietnam GR2 coffee samples while the *Flowery* note was from a Burundi coffee samples. Tables 1 and 2 report the sensory scores assigned by the panel.

Each sample was produced by a fresh cycle of roasting and grinding, starting from the same batch of green coffee beans. The roasting degree of each sample was carefully measured by ground bean light reflectance, with a single-beam Neuhaus Neotec Color Test II instrument (Genderkesee, Germany) at a wavelength of 900 nm on 25-30g of ground coffee. Roasting degree was set at 55°Nh, in order to be close to the international standardization protocol for cupping (SCAA, 2015). Samples were roasted within 24 hours prior to cupping and left for at least 8 hours to stabilize.

Table 1 Sensory Scores assigned to Vietnam coffee samples by the Lavazza's sensory panel

VIETNAM GR2					
Acid	Bitter	Body	Intensity	Overall Quality	Woody
0	5	8	8.5	8	7

Table 2 Sensory scores assigned to Burundi coffee samples by Lavazza's sensory panel

BURUNDI					
Acid	Bitter	Body	Intensity	Overall Quality	Flowery
4	0	8	8.5	9	8

2.2 Samples Preparation

According to the literature[10], 30g of coffee powder was extracted with 200mL of dichloromethane (DCM) under stirring conditions at ambient temperature for 1h. The extracts were then filtrated, and the volatiles were isolated by Solvent Assisted Flavour Evaporation (SAFE) distillation at 40°C.

The distillates were then separated into the acid and neutro-basic fractions; acidic compounds were removed by treatment with aqueous sodium hydrogen carbonate (0.5 mol/L; three portions of 200mL total volume) to obtain the neutral and basic volatiles (NBV). The alkaline aqueous phase was washed twice with DCM (50mL), acidified at pH=2 with HCl (6M) and extracted with DCM (3x50mL) to obtain the acidic volatiles (AV). AV and NBV were dried over anhydrous with Na_2SO_4 and each concentrated to 1 mL by a Vigreux column (50cm x 1 cm) followed by a microdistillation to reach a final volume of 200µL.

2.2 High-Resolution Gas Chromatography-Olfactometry (HRGC-O)

For HRGC-O, a gas chromatograph (8160-Fisons Instruments, Mainz, Germany) was used. Helium at a pressure of 80 kPa served as the carrier gas. Samples were injected by cold-on-column injection onto capillaries DB-5 or DB-FFAP column (both 30 m, 0.32 mm i.d., 0.25 µm film thickness; J&W Scientific, Agilent Technologies, Waldbronn, Germany). The end of the capillary was connected to a deactivated Y-shaped glass splitter (Chromatographie Handel Mueller, Fridolfing, Germany) dividing the effluent of the column into two equal parts, which were then transferred via two deactivated fused silica capillaries (50 cm × 0.25 mm) to a sniffing port and a flame ionization detector (FID), respectively. The sniffing port consisted of a cylindrically shaped aluminium device (80 mm length, 25 mm diameter) with a bevelled top and a central drill hole (2 mm) housing the capillary. It was mounted on one of the GC detector base and heated to 230 °C. The FID was operated at 250 °C with hydrogen (20 mL/min) and air (200 mL/min). Nitrogen (30 mL/min) was used as the make-up gas.
Injection volume: 1uL temperature program, from 40 °C (2 min) to 230 °C (5 min) at 6 °C/min.

2.3 Stable Isotopes Dilution Analysis (SIDA)

2,3-Pentandione, Furfurylthiol and 4-Ethylguaiacole were quantified by mean of the Stable Isotope Dilution Assay (SIDA).
The standard/analyte response factors (R_f) were determined by analysing mixtures of known amounts of the unlabelled target compound and the respective isotopically labelled internal standard in five different ratios (5:1, 3:1, 1:1, 1:3, and 1:5) on a GCxGC-TOF platform.
Table 3 reports the fragments selected as a Ti to quantify each odorant, the type of labelling and the parameters of the calibration (Response factors and R^2).

Table 3 SIDA compounds parameters

Odorant	Analyte m/z	Isotope label	Standard m/z	R_f	R^2 cal. curve
2,3 Pentandione	43	$^{13}C_2$	45	0.7062	0.9989
Furfurylthiol	114	$^{2}H_2$	116	0.7153	0.9977
4-ethylguaiacole	152	$^{2}H_4$	155-158	0.5576	0.9999

5g of coffee powder were weighted and suspended in 100mL of Dichloromethane (DCM). The suspension was spiked with different amounts of the labelled standard in order to add an amount of standard as close as possible to the concentration of the analyte in the matrix.
The mixtures were extracted at room temperature under stirring for 1h and then distilled by SAFE. The distillates were concentrate on a Vigreux columns and diluted 1:50 before being injected on a GCxGC-TOF-MS platform. Each sample was analyzed in two replicates; each replicate involved a complete new spiking, extraction and distillation protocol.

3. RESULTS AND DISCUSSION

3.1 AEDA, peak selection and identification

Table 4 and 5 list the smelling peaks perceived during the AEDA of both NBV and AV fractions of Vietnam and Burundi samples. More than 30 smelling peaks were detected in the Neutro-Basic Fractions while 10 peaks were detected in the Acid Volatiles fraction. Each peak is characterized by its Retention Index, the odour description and the Flavour Dilution Factor (FD) assigned to the peak in both samples during the AEDA. As expected there are not significant qualitative variations between the volatiles fractions of the samples; with the exception of peaks #13 and #22, all other peaks have been perceived in both samples.

The identification of highly volatile compounds (3-methylbutanal and 2,3-pentandione) was carried out on a static headspace HRGC-MS platform.

The system consisted of a Thermo Scientific Trace Ultra gas chromatograph (Dreieich, Germany) with a Chrompack purge-and-trap (PTI/TCT) injection system 4001 (Frankfurt, Germany) coupled to a Varian ion-trap mass spectrometer Saturn 2100 T (Darmstadt, Germany).

The headspace of different amounts of coffee powder (0.050 and 0.1g) were automatically sampled in different volumes (5000 and 500μL) by a Varian Combi Pal autosampler (Darmstadt, Germany) with a gastight syringe. After sampling, the volatile compounds were collected in a cryo trap cooled with liquid nitrogen at −150 °C. By rapidly heating the trap to 250 °C, the compounds were transferred (injected) onto the capillary column. The detection of the compounds was carried out by mass spectrometry, and the effluent was simultaneously sniffed using a Y-type glass splitter and two uncoated fused silica capillaries (50 cm × 0.3 mm i.d.).

Table 4 Important odorants perceived during the AEDA protocol in the NBV fractions of Vietnam (*Woody*) and Burundi (*Flowery*) samples.

NBV				
#	RI DB-FFAP	Odour Description	Vietnam	Burundi
1	943	Malty	128	32
2	980	Fat-Buttery	512	512
3	1020	Fresh-Berry	2	16
4	1063	Sweet-Fat	16	256
5	1095	Fresh-Berry	4	4
6	1240	Roasty-Sweet	8	4
7	1284	Solvent-Pungent	8	16
8	1293	Nutty-Sweet	64	4
9	1311	Roasty-Vegetables	16	64
10	1328	Fresh-Sweet	32	4
11	1337	Nutty	64	64
12	1395	Sweet-Fruity	256	64
13	1406	Brown-Cooked	512	-
14	1416	Roasty-Baked-Woody	64	8
15	1430	Earthy-Baked	256	512
16	1446	Green-Berry	512	32
17	1476	Malty-Roasty	256	1024
18	1495	Pungent-Solvent	512	32
19	1531	Pepper	512	1024
20	1550	Woody-Paper-Leather	64	16
21	1570	Sweet-Dry Fruit	16	32
22	1628	Roasty-Baked	256	-
23	1638	Sweet-Roasty-baked	32	128
24	1651	Roasty	8	32
25	1661	Cooked-Baked	512	32
26	1675	Cooked-Boiled	128	128
27	1831	Fruity	512	512
28	1878	Candy	512	1024
29	2004	Mushrooms	32	8
30	2057	Spicy-Cloves	16	256
31	2221	Cloves	512	1024
32	2420	Faecal	256	32
33	2472	Faecal	256	16

Table 5 Important odorants perceived during the AEDA of the AV fractions of Vietnam (*Woody*) and Burundi (*Flowery*) Samples.

AV				
#	RI DB-FFAP	Odour Description	Vietnam	Burundi
1	1672	Cheesy	512	512
2	1692	Berry	4	4
3	1718	Roasty	64	4
4	1805	Spicy-Roasty	16	64
5	1841	Spicy-Soup	256	256
6	1876	Sweet-Candy	1024	512
7	2052	Sweet-Baked	512	1024
8	2105	Solvent-Pungent-Leather	1024	128
9	2220	Cloves-Boiled Vegetables	1024	512
10	2409	Spicy-Pungent	64	32

In these tables, peaks with highest FD factors are the key aroma compounds of the investigated samples; the AEDA demonstrates their key role in the definition of the samples aroma.
Some peaks (e.g. #2, 19, 26, 27, 28, 31 of the NBV fractions and #1, 5, 6, 7, 9 in the AV fractions) not only have high FD factors but their behaviour is similar in the two samples under investigation. These odorants may be part of the "coffee chemical odour code"; in other words, they may be part of the pool of chemicals able to compose the general coffee aroma. This hypothesis is supported by the fact that their contribution to the aroma do not change even if the coffee samples under study have very different aromatic properties.
To find out the aromatic differences between these two samples in the Comparative AEDA peaks with high FD were not only considered but also all those peaks whose FD values differ between the two samples of at least two dilution points. This cut-off was chosen to minimize the errors in the peaks selection (figure 3).

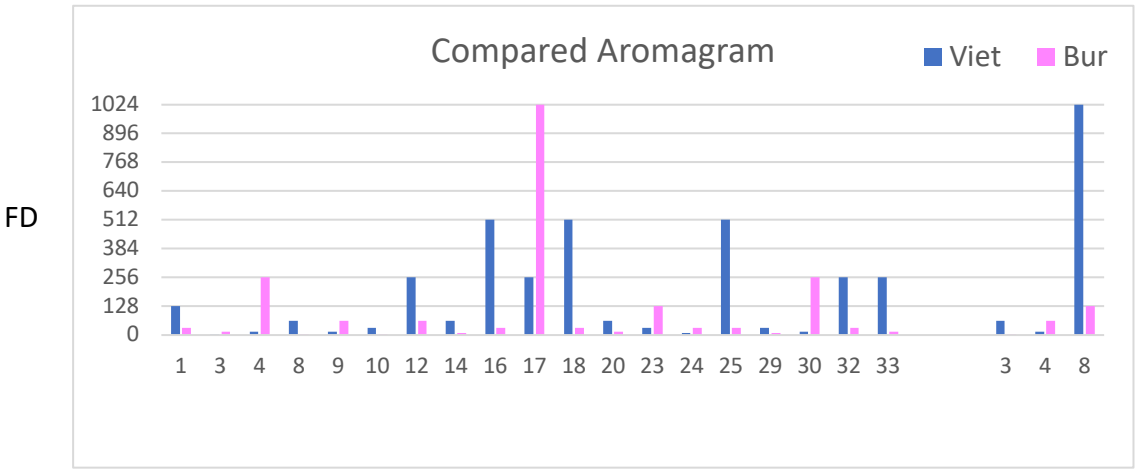


Figure 3 Compared Aromagram between Vietnam (Woody) and Burundi (Flowery) volatile fractions.Figure 3 show a visual representation of the FD factors of the 24 peaks selected as responsible for the aromatic differences between the two samples together with their FD factors.

Among this group, peaks characteristics of each sample elicits perception not always in agreement with our expectations: Vietnam samples was expected to be characterized by peaks with Roasty and brown notes. This expectation was satisfied by "exclusive Vietnam peaks" # 13 and 22 (brown

and Roasty notes) and by peaks #1, 21, 26, 33, 34 (responsible for Malty, Cooked-Baked, and Faecal notes) but not by peaks 13 and 17 which elicit Fruity and Berry odours. These observations can be confirmed also on Burundi characteristic peaks.

Table 6 Identification of Target peaks in the Neutro-Basic Fractions.

#	I^Ts	*Odour Description GC-O*	*Identification*	*CAS n°*	*Literature Odour description***	*Reference I^Ts*
1	943	Malty	3-methylbutanal	590-68-3	Malty	933
3	1020	Fresh-Berry	Ethylbutanoate **$**	105-54-4	Fruity-Juicy-Pineapple	-
4	1063	Sweet-Fat	2,3-Pentandione	600-14-6	Buttery	1058
8	1293	Nutty-Sweet	4.Methylthiazole	693-95-8	Nutty	1280
9	1311	Roasty-Veggie	2,5-dimethylpyrazine	123-32-0	Sweet-Green-Burnt	1314
10	1328	Fresh-Sweet	1-Hydroxy-2-butanone	5077-67-8	Sweet-Malty-Coffee	-
12	1395	Sweet-Fruity	2-ethyl-,5-methylpyrazine + 2-ethyl-6-methylpyrazine	13360-64-0 13925-03-6	Coffee-Nutty	1355-1359
13	1406	Brown-Cooked	2-Ethyl-3,5-dimethylpyrazine	15707-34-3	Burnt-Roasted-Popcorn	1410
15	1416	Roasty-Baked-Woody	Furfurylthiol	98-02-2	Sulphur-Roasty-Coffee	1391
17	1446	Green-Berry	2,3 Diethylpyrazine	15707-24-1	raw, nutty, green pepper	1396
18	1476	Malty-Roasty	2,3-diethyl-5-methylpyrazine	18138-04-0	Musty/Nutty/Potato	1481
19	1495	Pungent-Solvent	Acethylfuran	1192-62-7	Sweet-Almond	-
21	1550	Woody-Paper-Leather	Furfuryl Acetate	623-17-6	Green-Banana peel	1536
22	1620	Roasty-Baked	1-methyl-2-pyrrole carboxyaldehyde **$**	1192-58-1	Roasted-Nutty	-
24	1638	Sweet-Roasty-baked	1-Methylpyrrole-2-Carboxaldehyde	1192-58-1	Roasted-Nutty	-
25	1651	Roasty	2-Acethyl-1-methylpyrrole	932-16-1	Earthy-Nutty	1649
26	1661	Cooked-Baked	Furfuryl alcohol	98-00-0	Brown-Caramellic	1667
30	2004	Mushrooms	Difurfuryl ether**$**	4437-22-3	Coffee-Mushrooms like	-
31	2057	Spicy-Cloves	4-Ethyl-guaiacol	2785-89-9	Spicy-Smoky-Woody	2034
33	2420	Faecal	Indole	120-72-9	Faecal	2454
34	2472	Faecal	3-methyl-indole	83-34-1	Faecal	2504

** Literature odour description has been found on The Good Scents Company website: http://www.thegoodscentscompany.com

$ The identification still needs to be confirmed.

The identification of target peaks on both AV and NBV fractions has been obtained following the criteria described above; the results are reported in Tab 6 and 7.
The identification of compounds marked with **$** has been obtained only on the basis of their I^Ts, odour quality and MS spectra but was not possible to verify the identification with Pure Reference Compounds.

Table 7 Identification of target peaks on the AV fractions.

#	RI	*Odour Description GC-O*	*Identification*	*CAS n°*	*Literature Odour description***	*Reference RI*
1	1718	Roasty	3,4 dimethyl-2,5-furandione	766-39-2	-	1740
3	1805	Spicy-Roasty	3-methylcyclopentane-1,2-dione	765-70-8	Sweet-Coffee-Caramel	1830
4	2105	Solvent-Pungent-Leather	3-methylphenol	108-39-4	Woody-Leather-Phenolic	

** Literature odour description has been found on The Good Scents Company website: http://www.thegoodscentscompany.com

3.2 Quantitative information determination

2,3-Pentandione, Furfurylthiol and 4-Ethylguaiacole were quantified in both samples by mean of the Stable Isotope Dilution Assay (SIDA) in order to investigate if the concentrations were in agreement with the results of the AEDA and to obtain a further tool to compare the two approaches.

Table 8 Quantitation Results; green and red arrows indicate witch samples present the higher and the lower value of concentration for each compound.

Odorant	Vietnam (mg/L)	RSD%	Burundi (mg/L)	RSD%
2,3 Pentandione	▼ 2.03	5.60	▲ 9.40	0.75
Furfurylthiol	▲ 1.07	4.49	▼ 0.89	3.50
4-ethylguaiacole	▲ 8.54	1.24	▼ 2.05	5.09

Table 8 summarized the results of the quantitation performed by SIDA.

Another, simple but meaningful experiment has been performed on both the samples: an Internal Standard has been added to the suspensions before starting the work-up and then the targets normalized responses were compared between the samples.
More in detail: 30g of coffee powder were weighted and suspended in 200mL of DCM; 5µL of Ethylcycloexanoate from a 1000µg/L solution was added to the mixture. This compound has been chosen because it has a higher polarity compared to the correspondent alkane and better fits with the coffee composition.
Table 9 reports the targets normalized responses for both samples. Coloured arrows help to point out in which sample each compound has the higher normalized response.

Table 9 Normalized Responses obtained injecting the SAFE distillate in the GCxGC-TOF-MS platform

Odourant	Resp Norm Distillate	
	NBV	
	Vietnam	Burundi
2,3-pentanedione	▼ 3.050	▲ 27.210
4-methylthiazole	▲ 4.080	▼ 2.630
2,5-dimethylpyrazine	▼ 98.360	▲ 119.130
1-Hydroxy-2-butanone	▼ 1.360	▲ 8.830
Pyrazine 2-ethyl-, 5-methyl + Pyrazine 2-ethyl-, 6-methyl	▲ 30.190	▼ 5.770
2-Ethyl-3,5-dimethylpyrazine	▲ 18.680	▼ 4.850
Furfuryl thiol	▼ 0.950	▲ 1.000
2,3-diethylpyrazine	▲ 2.400	▼ 0.580
2,3-diethyl-5-methylpyrazine	▲ 4.220	▼ 0.810
Acethylfurane	▲ 27.830	▼ 21.940
Furfuryl Acetate	▼ 14.670	▲ 26.160
1-methylpyrrole-2-carboxyaldehyde	▼ 34.970	▲ 56.922
2-Acethyl-1-methylpyrrole	▼ 6.730	▲ 11.900
Furfuryl alcohol	▼ 188.170	▲ 213.330
Difurfuryl ether	▼ 11.530	▲ 19.250
4-Ethyl-guaiacole	▲ 22.110	▼ 3.350
Indole	▲ 8.070	▼ 2.420
3-methyl-indole	▲ 1.790	▼ 0.180
Odourant	**AV**	
	Vietnam	**Burundi**
3,4 dimethyl-2,5 furandione	▲ 8.7800	▼ 8.1200
1,2-Cyclopentanedione, 3-methyl-	▼ 1.5300	▲ 20.7800
3-methylphenol	▼ 2.4800	▲ 2.8000

3.3 Qualitative comparison between sensomic and sensometric target compounds.

The main goal is to compare the results obtained by applying two different approaches (sensomic and sensometric) in the definition of the chemical fingerprint of *Woody* and *Flowery* sensory notes.
The pool of compounds selected to describe the sensory notes have been qualitatively compared to understand if the information obtained about the samples with the two approaches is consistent. Eight compounds within the group identified until now after the cAEDA were found to be also part of the list used to develop the prediction models in the sensometric approach.

Table 10 Shared target compounds between Sensomic and Sensometric approach.

Woody Model	- 2,3-pentanedione - 4-Ethyl-guaiacol - Difurfuryl ether* - 2-Ethyl-3,5-dimethylpyrazine
Woody and Flowery Models	- 1-Hydroxy-2-butanone - Acethylfuran
Flowery Model	- Pyrazine 2-ethyl-, 5-methyl + 2-ethyl,6-methylpyrazine - 2,5-dimethylpyrazine

Table 10 reports compounds identified with the sensomic approach also present in the list of targets used to develop sensometric prediction models. The first box reports the compounds in common with the *Woody* Note Related Compounds (NRC) and the other two boxes the compounds in common with both *Woody* and *Flowery* NRC and with just *Flowery* NRC.

Another qualitative comparison between the two approaches can be done by observing of how aroma compounds pointed out with the molecular sensory science, behave within the chemical fingerprints obtained with the sensometric approach.
Table 11 considers three different compounds identified with molecular sensory science that apparently have not a key role in the *Woody* and *Flowery* models. The investigation of the correlation matrix obtained during sensometric data elaboration showed that several compounds used to develop the prediction models had high correlation coefficient with those identified at the DFA. This good relationship suggests that probably the huge differences between the two approaches both in the samples preparation and in the compounds selection criteria impact on the composition of the targets lists but not on the overall information that can be extracted from the sample.

Table 11 Note Related Compounds used in the Woody and Flowery prediction models highly correlated with Aroma compounds identified with the Sensomic approach. Pearson correlation coefficient was used

Key aroma sensomic		Sensometric Models compounds
2-Furanmethanethiol	highly correlate (r>0.75)	- 2,3-dimethyl Pyrazine, - Guaiacol - 1-methylethenylpyrazine - 4-vinylguaiacol - 3-ethyl-2,5-dimethylpyrazine, - 3,5-diethyl-2-methylpyrazine - phenolic derivate
Furfuryl alcohol	highly correlate (r>0.75)	- Maltol - Acetic acid - 2-Acetyl-5-methylfuran - Unk 17 - Ethenylpropanoate - Unk 14 - 5 Methyl Furfural - Acetylfuran - Furfural - 1-Hydroxy-2-butanone - difurfuryl ether - 2,3-Pentanedione
1-methyl-1H-Pyrrole-2-carboxaldehyde	higly correlate (r>0.70)	- 1H-Pyrrole-2-carboxaldehyde - 2-hydroxy-3-methyl-2-Cyclopenten-1-one - 3-ethyl-2-hydroxy-2-Cyclopenten-1-one - 2,3-Hexanedione - 5 Methyl Furfural - Acetoxyacetone - Acetylfuran - Furaneol - Furfuryl formate - Maltol - trans-Furfurylideneacetone - Unk 13b

3.4 Quantitative comparison between sensomic and sensometric target compounds.

Table 12 Comparison of quantitative data obtained with different strategies: AEDA, GCxGC-TOF normalized responses, SIDA quantitation, HS-SPME-GC-MS normalized responses. Similarly to the other tables green and red arrows helps the reader to individuate in witch of the samples each compounds is present in higher and lower amount.

Odourant		Resp Norm Distillate NBV		Quantification SIDA NBV		Norm Resp HS- SPME NBV			
	FD bur	Vietnam	Burundi	Vietnam (mg/L)	Burundi (mg/L)	Woody	no_Woody	Flowery	no_flowery
1-Hydroxy-2-butanone	4	1.360	8.830			0.004	0.014	0.014	0.007
2-Acethyl-1-methylpyrrole	32	6.730	11.900						
2,3-diethylpyrazine	32	2.400	0.580						
2,3-pentanedione	256	3.050	27.210	2.030	9.400	0.015	0.051	0.047	0.029
2,3-diethyl-5-methylpyrazine	1024	4.220	0.810					0.018	0.055
2,5-dimethylpyrazine	64	98.360	119.130			0.119	0.068		
3-methyl-indole	16	1.790	0.180						
4-Ethylguaiacole	256	22.110	3.350	8.540	2.050	0.254	0.040		
4-Methylthiazole	4	4.080	2.630						
Acethylfurane	32	27.830	21.940			0.121	0.321	0.314	0.186
Difurfuryl ether	8	11.530	19.250			0.014	0.066		
Furfuryl Acetate*	16	14.670	26.160						
Furfuryl alcohol	32	188.170	213.330						
Furfuryl thiol	8	0.950	1.000	1.070	0.890				
Indole	32	8.070	2.420						
2-ethyl-5-methylpyrazine + 2-ethyl-, 6-methylpyrazine	64	30.190	5.770					0.199	0.398
1-methylpyrrole-2-carboxyaldehyde	128	34.970	56.922					0.015	0.030

Odourant		AV		AV		AV			
	FD bur	Vietnam	Burundi	Vietnam (µg/L)	Burundi (µg/L)	Woody	no_Woody	Flowery	no_flowery
3,4 dimethyl-2,5 furandione	4	8.780	8.120						
1,2-Cyclopentanedione, 3-methyl-	64	1.530	20.780						
3-methylphenol	128	2.480	2.800						

Further comparisons were performed to investigate how each technique characterize the samples under study. Table 12 reports data on target compounds identified in this part of the project and compares the AEDA results with the normalized responses obtained in GCxGC-TOF, the quantitation data obtained by the SIDA and the normalized responses obtained by the HS-SPME-GC-MS of the coffee powder. The aim of this comparison is not to compare the numbers in their absolute values but to compare the trends that each compound has over the different techniques. Data on SPME were obtained by averaging the normalized responses among the group of samples selected as minimum and maximum expression of these sensory notes.
A good consistency between normalized data obtained from the HS-SPME and the direct injection of the SIDA distillate can be observed. This consistency suggests that even if the two analytical techniques are extremely different one another, the extracted overall information on the coffee samples volatile fractions is comparable. An example of this consistency is given by 2,3-pentandione: according to chemometric elaboration, it has been pointed out as a compound overexpressed in “Flowery” coffees; by comparing the SPME normalized responses, it is clear that it is higher in *Flowery* compared to *Woody* samples. This behaviour is confirmed by the comparison of normalized responses on the distillate and, above all, by the quantitation values obtained by SIDA.

When AEDA results are included in this comparison the scenario become more complex. Most compounds show a consistent behaviour between AEDA results and those obtained with the other strategies, but some others not. Moreover, some compounds have FD factors not in compliance with analytical data.
The reasons behind this behaviour might be related to a wrong identification: the identification procedure in molecular sensory science is complex because need to merge the smelling experience with the MS spectra and AEDA mandatorily required high operator training and a high number of replicates by GC-O to provide stable and robust results. This might be the case in particular for compounds like #10, 30, 21 whose identifications have not yet been confirmed by the Pure Reference Compound.

4. CONCLUSIONS

This part of the study aimed to compare sensometric and sensomic outcomes to “validate" the chemical fingerprints defined with the sensometric approach (section 2.2) Sensomic is nowadays considered the gold standard for food flavour investigations.

Promising data has been obtained by the comparison of the two techniques; a significant number of the compounds identified with the sensomic approach were also part of the selection used to develop the chemometric models (NRC) (Table 10) while those compounds identified in sensomic but not directly selected in the sensometric fingerprinting were highly correlated to them. This consistency confirmed the suitability of data elaboration workflow used in defining the chemical fingerprinting able to be used in the building of the prediction models of the sensory note. Another common point between the two strategies is the high correlation between some compounds identified with the sensomic approach and those used in the prediction models (Table 11).
In addition, most of these compounds show a good consistency between AEDA and the trend in the HS-SPME-GC-MS profiles (table 12). On the other hand, a lot of work has still to be done in order to increase the reliability of data obtained at the DFA and to investigate if and why there are some inconsistencies

Despite it is the reference approach for the determination of the foods key aroma compounds, the application of the sensomic approach for sensory notes profiling is extremely complex and time consuming since it would be necessary to run a comparative AEDA not only on two samples with high and low note scores but on a pool of these samples able to better represent the samples variability

In addition, this approach does not meet the need of developing an instrumental approach to predict samples sensory scores starting from chemical data.

From this preliminary comparison the results obtained by the sensometric approach seemed not far from those that may be obtained with the molecular sensory science, but in order to become representative of the panel evaluation, the study should be carried out for all sensory attributes defining the coffee aroma in order to obtain the chemical signatures of the coffee notes.

REFERENCES

1. Schieberle, P. & Grosch, W. Quantitative analysis of aroma compounds in wheat and rye bread crusts using a stable isotope dilution assay. *J. Agric. Food Chem.* **35,** 252–257 (1987).
2. Schieberle, P. & Grosch, W. Evaluation of the flavour of wheat and rye bread crusts by aroma extract dilution analysis. *Z. Lebensm. Unters. Forsch.* **185,** 111–113 (1987).
3. Grosch, W. Evaluation of the Key Odorants of Foods by Dilution Experiments , Aroma Models and Omission. *Chem. Senses* **26,** 533–545 (2001).
4. Pollner, G. & Schieberle, P. Characterization of the Key Odorants in Commercial Cold-Pressed Oils from Unpeeled and Peeled Rapeseeds by the Sensomics Approach. *J. Agric. Food Chem.* **64,** 627–636 (2016).
5. Kiefl, J., Pollner, G. & Schieberle, P. Sensomics analysis of key hazelnut odorants (Corylus avellana L. 'Tonda Gentile') using comprehensive two-dimensional gas chromatography in combination with time-of-flight mass spectrometry (GC×GC-TOF-MS). *J. Agric. Food Chem.* **61,** 5226–5235 (2013).
6. Steinhaus, M. & Schieberle, P. Characterization of Odorants Causing an Atypical Aroma in White Pepper Powder (Piper nigrum L .) Based on Quantitative Measurements and Orthonasal Breakthrough Thresholds AND. *Agric. Food Chem.* **53,** 6049–6055 (2005).
7. Engel, W., Bahr, W. & Schieberle, P. Solvent assisted flavour evaporation - a new and versatile technique for the careful and direct isolation of aroma compounds from complex food matrices. *Eur. Food Res. Technol.* **209,** 237–241 (1999).
8. Teranishi, R., Buttery, R. G., Stern, D.J. & Takeoka, G. Use Of Odor Thresholds In Aroma Research. *Food Sci. Technol. Technol.* **24,** 1–5 (1991).
9. Fallis, A. . & Lynch, S. Chemical Analysis of Food Techniques ans Application. Journal of Chemical Information and Modeling **53,** (2013).
10. Scheidig, C., Czerny, M. & Schieberle, P. Changes in key odorants of raw coffee beans during storage under defined conditions. *J. Agric. Food Chem.* **55,** 5768–5775 (2007).

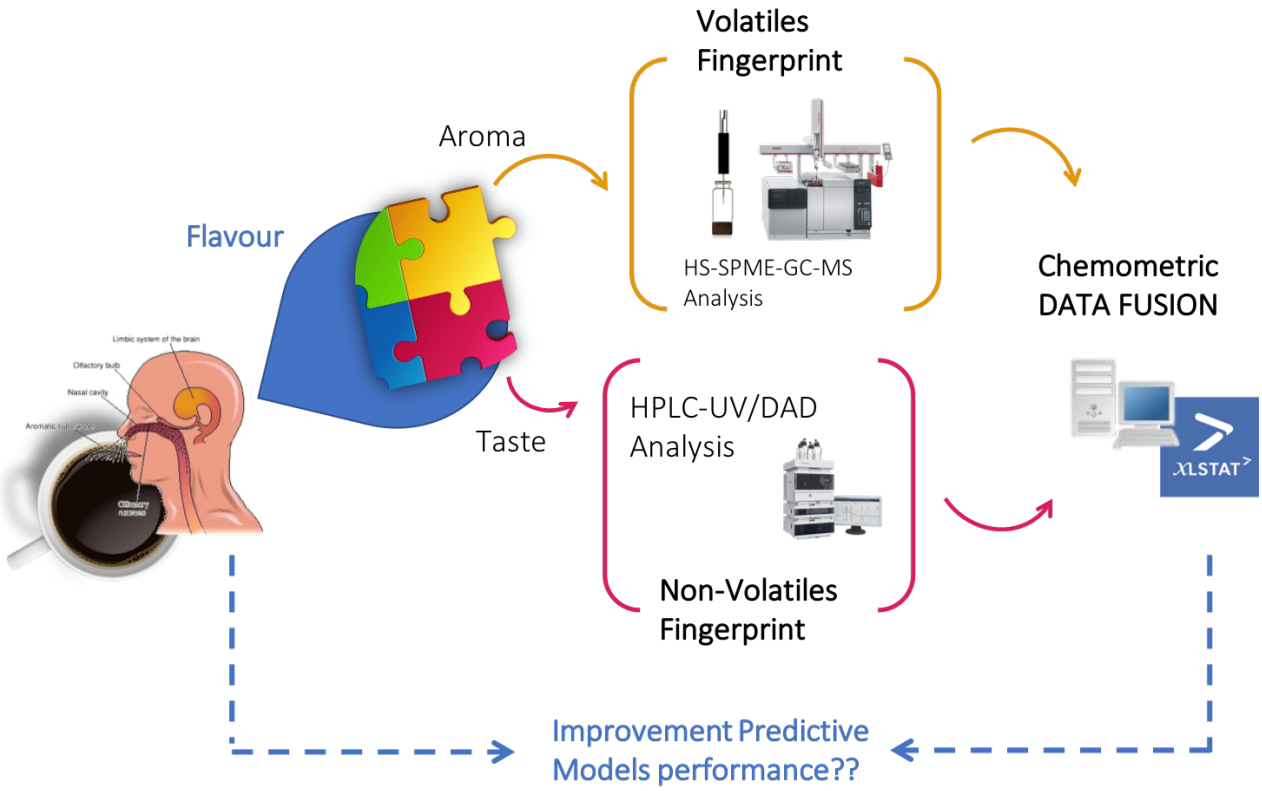
2.4 Multiple analytical platforms for coffee flavour investigations

INTERNATIONAL HUB
FOR COFFEE RESEARCH
AND INNOVATION

This work has been presented as an oral communication at the conference “The quality of coffee: a never-ending research”, Camerino, Italy November 30-December 1,2017
with the title:
“Sensory properties of a cup of coffee: possibilities and limits of a flavoromic approach for routine controls”
Liberto E.[a], Bressanello D.[a], Cordero C.[a], Sgorbini B. [a], Rubiolo P.[a], Pellegrino G.[b], Ruosi R.M.[b], Bicchi C.[a]

a Dipartimento di Scienza e Tecnologia del Farmaco, Università degli Studi di Torino, Via Pietro Giuria 9, Torino, Italy
b Lavazza S.p.A., Strada Settimo 410, Torino, Italy

0. GRAPHICAL ABSTRACT



1. INTRODUCTION

Coffee flavour is unique; its characteristics and its declinations as reported on the packaging of several coffee brands is of high appeal for consumers. Flavour perception is a complex feeling given by the interaction between the aroma, perceived through the sense of smell (orthosanally and retronasally) and the taste, perceived at the level of the tongue.[1] different classes of chemical compounds are responsible for these two feelings; the aroma is the result of the interaction of smell and volatile compounds with Odour Receptors (OR) while heavier and non-volatile compounds are related to the taste perception.

The evaluation of the coffee flavour is a key step of the whole coffee production chain from the selection of the row material to the creation of new valuable blends.

Although the scientific community has long studied since long time to overcome the contribution of the sensory panel their work is still indispensable[2]. The Specialty Coffee Association of America (SCAA) developed a standardized protocol for the coffee sensory analysis[3] to evaluate the quality of the incoming beans to be used in specialty mono origin coffees or in blends. SCCA protocols are also known as "Brazilian method" and are internationally used to evaluate the coffee beans quality and to define the price before further treatments. Evaluation is done on light roasted coffee samples to a specific colour (55–60° Nh), after eight hours of stabilization. The coffee sensory properties are evaluated by sniffing the powder and then the brew obtained by filter method. In the following, coffee samples are tasted by spraying the brew in the month. This multistep protocol allows the panelists to evaluate different coffee attributes; some of them more linked to the aroma (sensory notes like *flowery*, *fruity*, *woody*, *nutty* and *spicy*) and other more related to the tasting experience (*acidity* and *bitterness*). The quality and intensity of each attribute are evaluated simultaneously by using a scale varying from 1 to 10.

Several analytical approaches have been developed over the years to try to understand the real impact of the chemical composition on the food flavour profiles.

Pioneers of this branch of analytical chemistry were Czerny and Grosch with their studies [4,5] that led to the definition of molecular sensory science given by Schieberle et al in 2011[6,7].

This approach is undoubtedly effective to study aromatic food properties and affords to develop synthetic recombinant with very close sensory properties to those of the food matrix object of the study, but has some limitations, in particular when heavy components (related to the taste) are included in the overall food evaluation.

Flavoromics is an alternative approach, introduced from 2008 by Reineccius et al.[8], in which the metabolomics and, more in general, the -omics concepts are applied to the flavour and fragrance field. Flavoromics try to answer to some issues that were neglected by more conventional approaches such as the multimodal nature of flavour and the interaction between aroma and taste perceptions. This approach is less biased by the compounds selection and gives an overall vision of the secondary metabolome that can rule the sensory expression.

Flavoromics is an analytical untargeted methodology that studied the relationship between the chemical composition of foods and their sensory impact using chemometric tools. Compared to conventional approaches, its novelty lies on two main aspects: i) all potential sensory active compounds are kept in consideration without focusing just on known key flavour compounds, and ii) this approach is data-driven and each compound is considered not as a function of its chemical classification but just according to the distribution of its analytical output through the samples under study.[8]

This approach affords measuring more stimuli in an attempt to link analytical data to perception and potentially gives a better prediction of Flavour since it includes inputs from more chemical compounds as well as identifying new Flavour contributors.

At this stage, two analytical platforms combined with the sensory descriptive analysis, were developed and compared with the aim to simulate the main phases of the SCAA cupping protocol (smell and taste). The results of HS-SPME-GC-MS of the powder selected in previous study as the most satisfactory sampling technique for routine screening in terms of informative level, speediness and possibility of automation have been combined with the HPLC-UV/DAD analysis of the brews.
The coffee brew has been prepared according to the SCAA protocol and non-volatiles chemical fingerprints were used in data fusion with those obtained from the HS-SPME-GC-MS and HPLC-UV/DAD analyses and with the sensory descriptive results to evaluate the ability to describe the sensory quality of coffees. Aroma and taste evaluation aimed to understand how the different data sources are connected one another and their comparison would be useful, within its limitations, to understand if the instrumental data are of help to define the multimodal perception.

2. MATERIALS AND METHODS

2.1 Reagents and matrices

Coffees samples, consisting of roasted coffee ground to suit a coffee-filter machine, were kindly supplied over a period of 12 months by Lavazza Spa (Turin, Italy). 155 coffee samples with distinctive sensory notes originating from different countries belonging the *Coffea Arabica* L. (Arabica) and *Coffea canephora* Pierre (Robusta) were analyzed. The roasting degree of each sample was carefully measured by ground bean light reflectance, with a single-beam Neuhaus Neotec Color Test II instrument (Genderkesee, Germany) at a wavelength of 900 nm on 25-30g of ground coffee. Roasting degree was set at 55°Nh, in order to be close to the international standardization protocol for cupping. Samples were roasted within 24 hours prior to cupping and left for at least 8 hours to stabilize.
The coffee brew for cupping and analysis was prepared from 18g of coffee powder and 300mL of water, using a commercially available “Xlong” coffee filter machine. Two milliliters of brew are then filtered on a 0.2 µm 13 mm nylon membrane syringe filters (Agilent, Little Falls, DE, USA) and 20 µL directly injected.

2.2 Descriptive sensory analysis of coffee samples

The samples have been submitted to sensory evaluation through quantitative descriptive analysis by a panel of trained coffee experts following the SCAA protocol[3].
The protocol includes three tasting steps: i) evaluation of the aroma by sniffing the dry ground coffee, ii) evaluation of the aroma by sniffing the brew three minutes after its preparation and stirring, and iii) evaluation of the Flavour after 8-10 minutes. Attributes such as aftertaste, acidity, body, and balance are evaluated by tasting the brew, spraying it into the mouth to maximize retro-nasal vapors. Cup quality was assessed for several flavour attributes such as: *Acid*, *Bitter*, *Flowery*, *Fruity*, *Spicy, Nutty* and *Woody*. The intensities of each attribute have been evaluated simultaneously, upon a scale from 0 to 10.

2.3 Volatile fraction analysis

The volatile fraction of the samples responsible for the aroma has been analyzed by sampling the headspace of the dry grounded coffee powder by Solid Phase Microextraction (SPME) followed by GC-MS analysis. Sampling conditions, instrumental set-up and chemicals identification are reported in section 2.2.

2.4 Non-volatile Fraction analysis

Data on the non-volatile fraction have been acquired from a HPLC-UV/DAD Agilent 1200 system provided with a Spectra System UV Diode Array Detector 1100 series (Agilent, Little Falls, DE, USA). Data acquisition and data handling were performed with Chemstation LC 3D system software Rev 3.03 01-SR1 (Agilent, Little Falls, USA). HPLC column was a Platinum EPS C18 (250×4.6 mm, 80A, 4 µm) (Alltech, Deerfield, USA).

Operative conditions: injection volume, 20 µL; detection wavelength, 325 nm for cinnamic and chlorogenic acids (monomers and dimers derivatives) and 276 nm for caffeine and trigonelline; mobile phase, ACN/H2O–ACN 0.1% formic acid; flow rate, 1.0 mL/min; mobile phase program: from 85% H2O (7 min) to 45% H2O (20 min) to 100% ACN hold for 2 min. Before re-injection, the HPLC system was stabilized for at least 5 min.

Compounds identification have been carried out by injecting the brew (5 µl) in a Shimadzu Nexera X2 system equipped with a photodiode detector SPD-M20A in series to a triple quadrupole Shimadzu LCMS-8040 system provided with electrospray ionization (ESI) source (Shimadzu, Dusseldorf Germany).

Samples were analyzed on an Ascentis Express C18 column (15 cm x 2.1 mm, 2.7µm, Supelco, Bellefonte, USA) using water/formic acid (999:1, v/v) and acetonitrile/formic acid (999:1, v/v) as mobile phases A and B, respectively. The flow rate was 0.4 mL/min and the column temperature was maintained at 30°C. The gradient program was as follows: 15% B for 7 min, 15-55% B in 3 min, 55-100% B in 1.5 min, 100% B for 1 min. Total pre-running and post-running time was 23 min. MS operative conditions were as follows: heat block temperature: 200°C; desolvation line (DL) temperature: 250°C; nebulizer gas flow rate: 3 L/min drying gas flow rate: 15 L/min. Mass spectra were acquired both in positive and in negative full-scan mode over the range 100-1000 m/z, event time 0.5 sec. Product Ion Scan mode (collision energy: - 35.0 V for ESI+ and 35.0 V for ESI-, event time: 0.2 sec) was applied to compounds for which a correspondence between the pseudomolecular ions [M+H]+ in ESI+ and [M-H]- in ESI- had been confirmed. Multiple Reaction Monitoring acquisition (collision energy: - 35.0 V for ESI+ and 35.0 V for ESI-, dwell time: 20) was carried out on specific product ions derived from precursor ions fragmentation.

Chemicals:

HPLC-grade acetonitrile (LC-MS grade), formic acid (>98% purity), de-ionized water (18.2 MΩ cm) was obtained from a Milli-Q purification system (Millipore, Bedford, MA, USA).

Cryptochlorogenic acid, 3,5-dicaffeoylquinic acid and 4,5-dicaffeoylquinic acid were from Phytolab (Vestenbergsgreuth, Germany). Chlorogenic acid, Neochlorogenic Acid, 3,4-dicaffeoylquinic acid, Trigonellin and Caffein were from Sigma Aldrich (Bellefonte, USA).

2.5 Data Processing

Data (Sensory, GC-MS and HPLC) were explored by PCA followed by Multiple Factorial Analysis (MFA). The latter chemometric tool allows to compare and to investigate the relationship between different data matrices.

Features selection for each analytical platform related to the sensory note was made by Partial Least Square-Discriminant Analysis (PLS-DA) between samples able to minimize and maximize the sensory note expression. This procedure has been used to define features used afterword, singularly (Aroma or Taste) or comprehensively (“data fused”, i.e. both data sources together), to develop PLS prediction models of the sensory scores. All data elaborations have been performed on XLSTAT (version 2015.5.01.23.164) copyright Addinsoft 1999-2015.

3. RESULTS AND DISCUSSION.

First data exploration has been carried out by a Principal Component Analysis (PCA) on both Aroma and Taste data matrices in which each sample (observation) is described by the pool of detected features (variables).
In this elaboration, Aroma and Taste data have been kept separate in order to see the samples behavior within the single approach and therefore the role of the chemical information encrypted in each analytical platform in describing the samples.

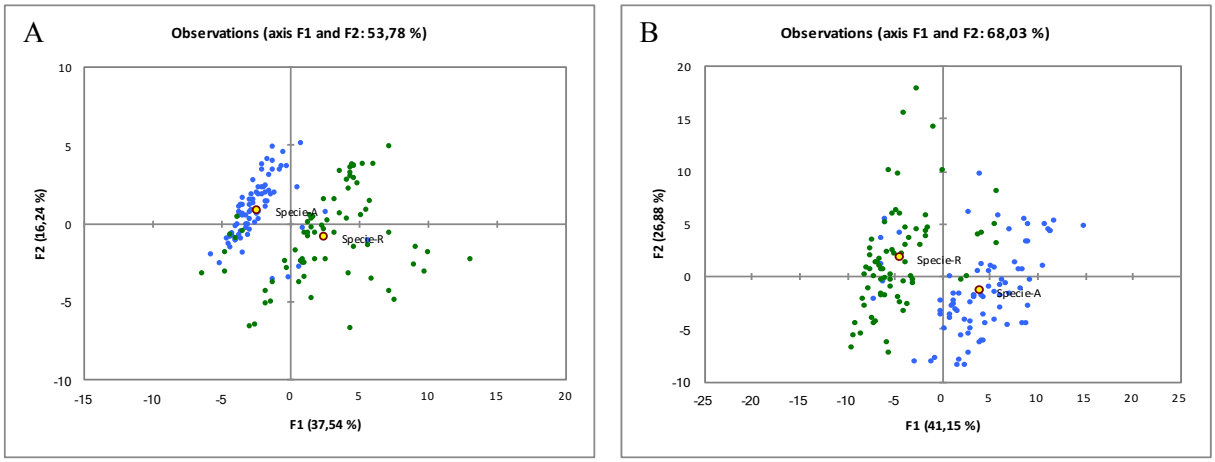


Figure 1 Reports the scores plots obtained from the PCA of Taste (A) and Aroma (B) data respectively

Figure 1 shows the samples distribution on the two PCs able to cover more than the 50% of the total variance of the data matrix (53% in the case of Taste (fig.1 A) and 65% for Aroma) (fig.1B)). In both diagrams, a good separation between Arabica (Blue) and Robusta (Green) samples can be observed. This consistency of the two different data sets suggests that volatile and non-volatile fractions give a similar contribution, from a chemical point of view, to the samples chemical characterization.

Multiple Factor Analysis (MFA) has then been applied to compare the three different data matrices (GC-MS and HPLC-UV/DAD fingerprints, and descriptive sensory analysis).
For sensory evaluation all the seven sensory notes (*Acid*, *Bitter*, *Flowery*, *Fruity*, *Nutty*, *Woody* and *Spicy*) have been considered.
MFA proceeds in two steps: First it computes a PCA of each data table and 'normalizes' each data table by dividing all its elements by the first singular value obtained from its PCA. Second, all the normalized data tables are aggregated into a grand data table that is analyzed via a (non-normalized) PCA that gives a set of factor scores for the observations and loadings for the variables[9]. Results, expressed as RV coefficients, indicate to what extent the tables/variables distribution are related two by two; The more the variable are related one another, the higher is the RV coefficient (variation range 0-1) (Table 1)

Table 1 MFA RV coefficients

	Sensory	Taste	Aroma	AFM
Sensory	1.0000	0.4708	0.5923	0.8229
Taste	0.4708	1.0000	0.5066	0.7987
Aroma	0.5923	0.5066	1.0000	0.8557

AFM	0.8229	0.7987	0.8557	1.0000

In this case the mutual correlation between Aroma and Taste was 0.5066, between Sensory data and Taste was 0.4708 and, Sensory data and Aroma was 0.5923 (Table 1).
These data suggest a certain relationship between the different data matrices even if the correlation is not that strong. This elaboration confirms what already reported in the literature, i.e. i) the Aroma has an important role in the definition of the sensory profile and ii) the highest correlation coefficient is between aroma and sensory data.
The reported correlation values (AFM values) suggest that the multicollinearity between the information provided by the two approaches is weak; this is important because means that they both contribute to the definition of the overall coffee flavour and that any aspect can totally be neglected.

Next elaborations have been performed to investigate how Aroma and Taste data fusion behaves in the modelling of the six different sensory notes under study compared to the Aroma and Taste data considered singularly.
The aims of these elaborations were:

- to evaluate the impact of HPLC data of the model performance when merged with HS-SPME-GC-MS data
- to evaluate the predictive power of the non-volatile fraction and its suitability in the development of sensory score prediction models compared to the volatile fraction

Typically, *bitter* and *acid* notes are perceived as taste attributes, the bitter note was therefore studied to see if with the non-volatile fraction was correctly described in terms of score prediction compared to volatiles fraction, and if the data fusion afforded more powerful or better information in its description.

3.1 *Bitter* flavour evaluation

Data sets (GC and HPLC) have been initially processed separately and then fused into a single one. For both analytical platforms the features highly related with the expression of the *Bitter* note have been selected by Partial Least Square-Discriminant Analysis (PLS-DA).
The Variable Impact on Projections (VIP) values have been used as selection parameter; the cut-off has been chosen by plotting all VIPs on a histogram and looking at the point at which the VIP values dropped sharply.
Variable importance in projection (VIP) coefficients reflect the relative importance of each X variable in prediction or classification models.
The prediction model has been developed by applying a Partial Least Square regression algorithm on the selection of *Bitter*-related features

155 samples were considered to build the regression model: among them 20 have been randomly chosen and used as a validation set, while 30 where excluded from the training set and used as an external test set.
The models (HS-SPME-GC-MS, HPLC-UV/DAD and data fusion) have been evaluated and compared on the basis of their model quality index (Q^2), their Coefficient of Determination (R^2) and the Root Mean Squared Error Cross-Validation (RMSECV) and Prediction (RMSEP).
These parameters have been described in section 1.3.
Q^2 and R^2 values ranges between 0 and 1 and the closer they are to 1 the better it is, while RMSEs (CV and P) operates on the opposite, since they are expressions of an error the lower they are, the better it is.

Table 2 Bitter PLS model's parameters

Bitter	*Aroma*	*Taste*	*Aroma +Taste*
n° variables	22	14	39
Q^2	0.742	0.666	0.692
R^2	0.892	0.810	0.888
RMSECV	0.579	0.659	0.575
RMSEP	1.073	0.929	1.120

The observation of the values reported in table 2 unexpectedly suggests a similar behavior of the data from Aroma and from the fused matrix (Aroma +Taste) in the description of the *bitter* note. The volatile fraction (evaluated by HS-SPME-GC-MS (Aroma)) shows better performance in the chemical description of the expression of the *bitter* note, the with highest values of both Q^2 and coefficient of determination (R^2).
On the other hand, the non-volatile fraction, shows a lower ability in the description of the *bitter* note. Although this analysis applies target wavelengths characteristic of bitter-related chemicals, (i.e. Caffeine, Trigonelline and Chlorogenic Acid derivatives), the regression parameters are not so good.
This can probably be due to other inferences on the description of this note that are not detectable within these wavelengths. The volatile fraction provides useful information to characterize the evolution of this note within the pool of samples.

The investigation of *Bitter-related* aroma compounds afford the identification of several pyrazines: (2-isopropenylpyrazine, 2-*n*-propyl pyrazine, 1-methylethenylpyrazine, 2,3-dimethylpyrazine, 2-methyl-6-(1-propenyl)pyrazine, 2,3-diethyl-5-methylpyrazine§, 2-ethyl-3,5-dimethylpyrazine§), guaiacoles (4-ethylguaiacol§, guaiacol§, 4-vinylguaiacol§), phenethyl alcohol, acetic acid, furfural, 1-hydroxy-2-butanone, 1H-pyrrole-2-carboxaldehyde, 2-furanmethanethiol§, furfurylmethylsulfide, 2,3-pentandione§ and 2,3-butandione§.
Within this selection many compounds (§) are in the list of the coffee key-odorants defined by Blank et al. through a Molecular Sensory Science approach[10]. It's interesting that these compounds are described by earthy, roasty, burn and phenolic notes and none of them is directly related to Bitterness. The close relationship highlighted by the PLS algorithm between these compounds and the evolution of the sensory note open new interesting perspectives in the study of food flavour that looks beyond the conventional approaches.

The same elaboration has been made with the non-volatile fraction; the model performance reported in table 2, as expected, confirms the close relationship between the evolution of the *bitter* note and the pool of compounds analyzed in HPLC. Bitter related features on the chromatographic fingerprints were identified (or tentatively identified) through a retrospective analysis by HPLC-DAD-MS (figure 2).

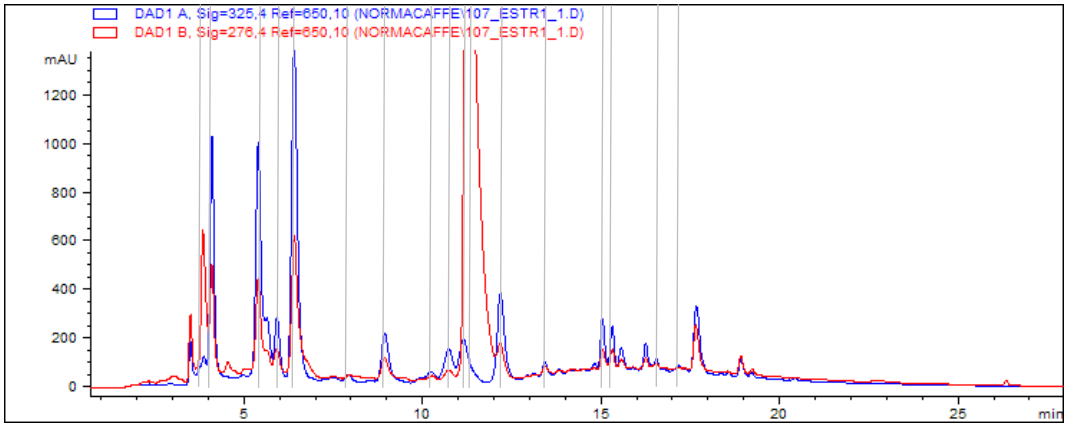


Figure 2 HPLC coffee extract fingerprint; grey bars show features related to the bitter note and subjected to a deep investigation by HPLC-DAD-MS.

Some of the main components were identified by comparing their retention times, UV and MS spectra to those of authentic standards while other components were tentatively identified on the basis of their UV and mass spectral information, compared to those reported in the literature[11,12]. (table 3) MS analysis were run by monitoring the multiple reaction (MRM) transitions of precursor ions (m/z)).

Three different feruloylquinic acid (FQA) isomers, 3-caffeoylquinic acid (CQA) and 5-CQA[§], n=2 caffeoylquinic lactone (CQL) isomers[§], 3,4 and 3,5 diCQA, caffeine[§] and trigonelline[§] have been identified among the *bitter* related features.

Similarly to those in the volatile fraction, also some of these non-volatile compounds (§) have been previously associated to the *bitter* note by studied of molecular sensory science performed by Hofmann and his group [13–15].

The potential gained in explanatory and predictive power by combining chemical information (i.e., relative to volatiles and non-volatiles) collected from different instruments on the same set of samples has also been investigated. Since the flavour perception derives from the interaction between aroma and taste, the combination of the information provided by different fractions was expected to improve the predictive model performance.

The data fusion model's performance (table 2) partially disregarded this expectations since the combined model show performance in line with those registered with the single Aroma and Taste models. Data fusion did not improve either the error committed on the cross-validation set (RMSECV), or the prediction of the external test set (RMSEP).

Table 3 *List of identified and putatively-identified compounds in the coffee brews. Underlined compounds have been identified with the pure standard while the hypothesized identification have been done comparing compound features with those present in the literature [11,12]. Each compound is quoted through its relative retention time, UV spectrum, pseudomolecular ions and molecular weight fragments obtained by Product Ion Scan mode (PIS).*

Compound Name	RT (min)	λmax (nm)	Mol. weight (g/mol)	[M–H]+ m/z	[M–H]– m/z	MS^2 + m/z	MS^2 - m/z
Trigonellin	0.769	263	137	138			
5-O-Caffeoylquinic acid	1.122	323/233	354	355	353		
3-O-Caffeoylquinic acid	1.403	322/239	354	355	353		
4-O-Caffeoylquinic acid	1.474	324/235	354	355	353		
Caffein	1.588	270	194	195			
3.4 dicaffeoylquinic acid	8.136	324	516	517	515		
3.5 dicaffeoylquinic acid	8.973	322/299	516	517	515		
4.5 dicaffeoylquinic acid	9.637	325	516	517	515		
Caffeoylquinic acid lactone	1.971	325/296	336	337	335	163	161/133
Caffeoylquinic acid lactone	2.894	326	336	337	335	163/117	161/133
Caffeoylquinic acid lactone	3.188	326	336	337	335	163/117	161/133
Caffeoylquinic acid lactone	3.475	326	336	337	335	163/117	161/133
feruloylquinic acid	1.588	323	368	369	367	177/149/145/117	134/149/191
feruloylquinic acid	2.14	310	368	369	367	177/149/145/117	134/149/191
coumaroylquinic acid	2.14	310	338	339	337	147/119	173/119
feruloylquinic acid	2.549	323	368	369	367	177/149/145/117	134/149/191

3.2 Flavoromic investigation of all considered sensory notes

The same data elaboration workflow has been applied to investigate the ability of the different data sources to describe the other flavour attributes with the aim to investigate if the non-volatile fraction could have a role in their description.

Each sensory note described by the panel has been modelled using volatile, non-volatile and data fusion fingerprints; Table 4 reports a summary of the model performance.

Table 4 Summary of the PLS models' performance for the six investigated notes.

PLS model performance			
Acid	Aroma	Taste	Aroma + Taste
n° variables	22	10	26
Q^2	0.723	0.450	0.703
R^2	0.829	0.636	0.825
RMSECV	0.594	0.854	0.605
RMSEP	0.898	1.069	0.875
Flowery	Aroma	Taste	Aroma + Taste
n° variables	20	14	27
Q^2	0.223	0.199	0.099
R^2	0.585	0.498	0.597
RMSECV	0.806	1.042	0.847
RMSEP	0.972	1.067	1.020
Fruity	Aroma	Taste	Aroma + Taste
n° variables	19	16	39
Q^2	0.158	0.184	0.033
R^2	0.607	0.508	0.786
RMSECV	0.814	0.922	0.619
RMSEP	0.615	0.610	0.876
Spicy	Aroma	Taste	Aroma + Taste
n° variables	22	16	32
Q2	0.320	0.331	0.458
R2	0.709	0.720	0.823
RMSECV	1.063	1.051	0.821
RMSEP	1.066	0.971	1.217
Woody	Aroma	Taste	Aroma + Taste
n° variables	23	9	37
Q2	0.708	0.472	0.706
R2	0.879	0.714	0.885
RMSECV	0.798	1.228	0.782
RMSEP	0.920	0.948	1.129

For a clearer comparison, performance values of Taste models have been reported in the histogram below (figure 3).

From the Q^2 values (green bars); it is clear that the non-volatile fraction does not impact equally on the five sensory notes: *Acid*, *Spicy* and *Woody* notes show better performance compared to *Flowery* and *Fruity* notes. This trend is almost confirmed by the R^2 that are higher in *Acid*, *Spicy* and *Woody* notes.

The RMSECV values behave in a slightly different way: these values are in compliance with the previous observations on *Bitter* in the case of *Acid* note while the trend is not that clear for the other notes. A lower impact of the non-volatile fraction on *Fruity* and *Flowery* notes was expected since these notes are considered typical aroma attributes.

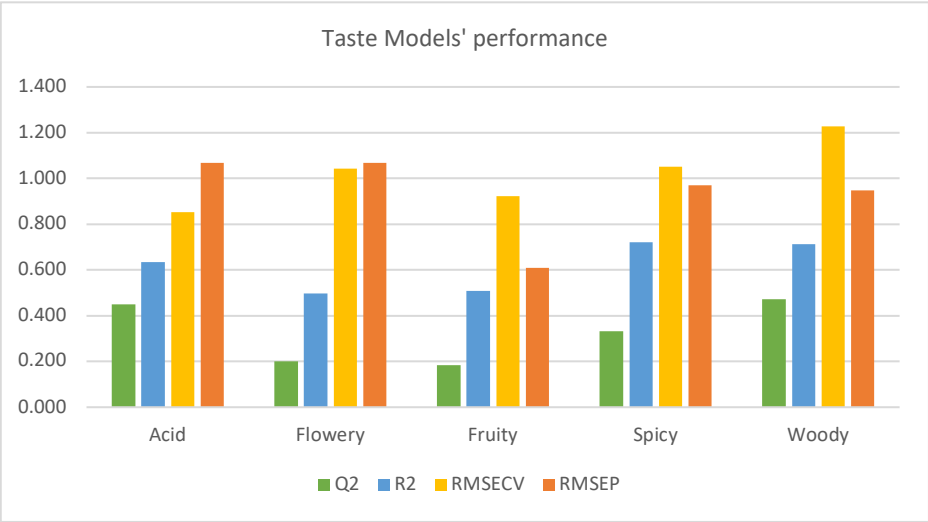


Figure 3 summary of PLS model performances on HPLC data

A better comparison between the performance of the models built up with Aroma and Taste data singularly or by combining them together (data fusion) is shown in the spider diagrams reported in Figure 4.

The first comparison is between Taste and Aroma models: most of the notes (with the exception of the *Spicy* note) show Aroma model performance (green lines) better than Taste ones. This observation suggests a better agreement with the HS-SPME-GC-MS data on the variation of the sensory scores in the samples set. The *Spicy* note partially agrees with the other notes. Aroma and Taste model performance are approximately superimposable.

From the performance from merging Aroma and Taste data, three different scenarios can be observed:

- *Acid* and *Woody* notes data fusion models (yellow lines) show an acceptable performance (Q^2 around 0.7, R^2>0.8 and a RMSECV lower than 1). Their overlapping with those obtained with PLS models only with Aroma data suggests that the inclusion of the Taste data provides a negligible improvement in the Aroma predictive models
- the second scenario involves *Flowery* and *Fruity* notes; the performance on the combined models (yellow lines) are lower than those obtained only with Aroma data.
  As expected, this worsening suggests that the volatile fraction is the best data source to investigate the expression of these sensory notes; the inclusion of Taste data not only does not provide any improvement (since its performance are lower (Figure 3)), but also it increases the noise around the information.
- The third scenario concerns the *Spicy* note: the performance achieved by the combined model are slightly better than those obtained with the models built only with Aroma and Taste data alone. The most significant improvements can be observed in the values of Q^2 and RMSECV (Table 3).

The performance improvement obtained combining Taste and Aroma compounds is supported by the fact that some key spicy volatile compounds (mainly phenolic compounds like Guaiacoles) originated from the thermal degradation of chlorogenic acids (exactly those observed in HPLC analyses)[16].

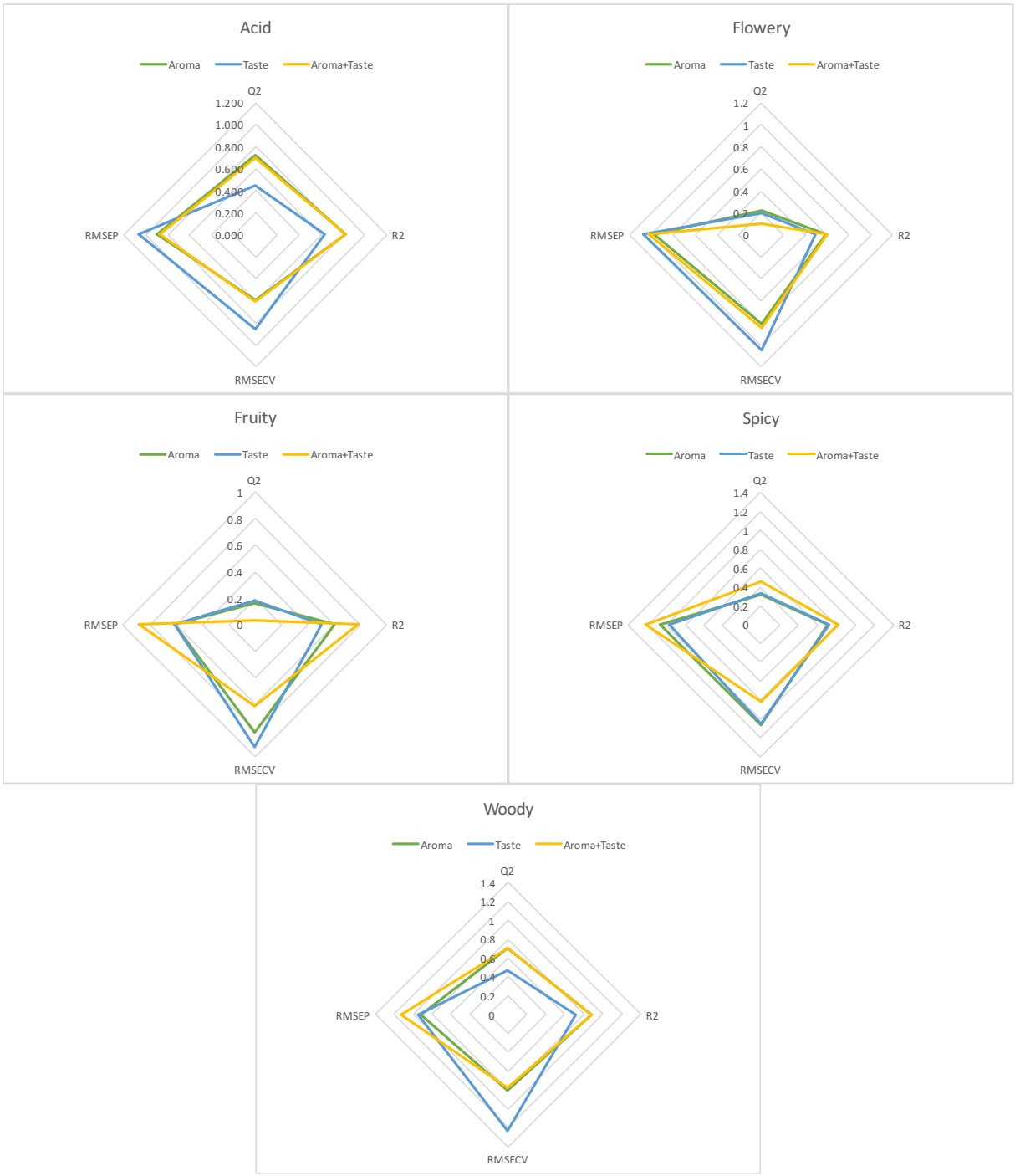


Figure 4 Spider diagrams reporting the models' performance registered in the regression models of each sensory note.

4. CONCLUSIONS

Coffee samples were analyzed according two different strategies to investigate two different flavour characteristics of this complex matrix. The volatile fraction was analyzed through HS-SPME-GC-MS of the roasted powder while the non-volatile fraction with HPLC-UV/DAD of the coffee brew prepared following the SCAA cupping protocol. The chemical information obtained from the analyses were related to those obtained by the descriptive sensory analysis performed by a trained panel.

One major barrier to accurate predictions of flavour is given by the multi-dimensionality of the chemical stimuli involved in Flavour perception. The Flavoromic approach, considering many compounds in the sensory notes modelling, takes care of possible correlations within and across the stimuli, generally not accounted by traditional research methods.

Data, compared by chemometric tools, support this hypothesis; the samples distribution after a PCA elaboration (Fig 1) confirmed that the sample characteristics were kept even though the analytical approach was very different. The samples distribution was comparable between aroma and taste analyses.

The MFA analysis confirmed a certain agreement between the three data matrices but the relatively low correlation values suggest that the information provided by the two analytical techniques are not redundant, and therefore they both contribute to the flavour definition, represented by the sensory matrix.

The contribution of both volatile and non-volatile fractions (as such and in combination) to the sensory scores prediction models were also evaluated.
The model performance suggests a different impact of the non-volatile fraction on the chemical fingerprints of the six sensory notes. As expected *Fruity* and *Flowery* notes are the less affected by the composition of the non-volatile fraction. Conversely, for *Bitter*, *Acid* and *Woody* notes, the non-volatile fraction gives a contribution to the samples characterization, although the improvement it produces on the performance of the regression model in the sensory scores prediction it is rather negligible.

Though MFA suggests a certain orthogonality between Aroma and Taste data (Table1), the PLS model performance highlights the key role of the volatile fraction, and therefore of the Aroma, in the sample sensory characterization. The performance of the PLS models built up from HS-SPME-GC-MS data (Table 2 and 4) are comparable to those obtained from the data fusion. Some authors report[1] that the flavour perception in all its aspects is mostly linked to the aroma composition and impact[17,1,18]. They assert, that, since the taste experience is mediated by a limited number of receptors and can only discriminate five different perceptions (acid, bitter, sweet, salty and umami), it has to be considered as a minor sense, whereas the sense of smell, being mediated by a higher number of different receptors, is a major sense. These limitations of the sense of taste are confirmed by the poor (or negligible) performance improvement registered in the data fusion models.

These observations together with the good performance obtained when the aroma data were used to define *Acid* and *Bitter* notes (considered as “typical taste notes”) chemical fingerprints, make it possible to hypothesize that the aroma analysis by HS-SPME-GC-MS provides an accurate and reliable representation of the food (coffee in our case) flavour, as it is perceived by a trained panel during a validated and standardized descriptive sensory analysis.

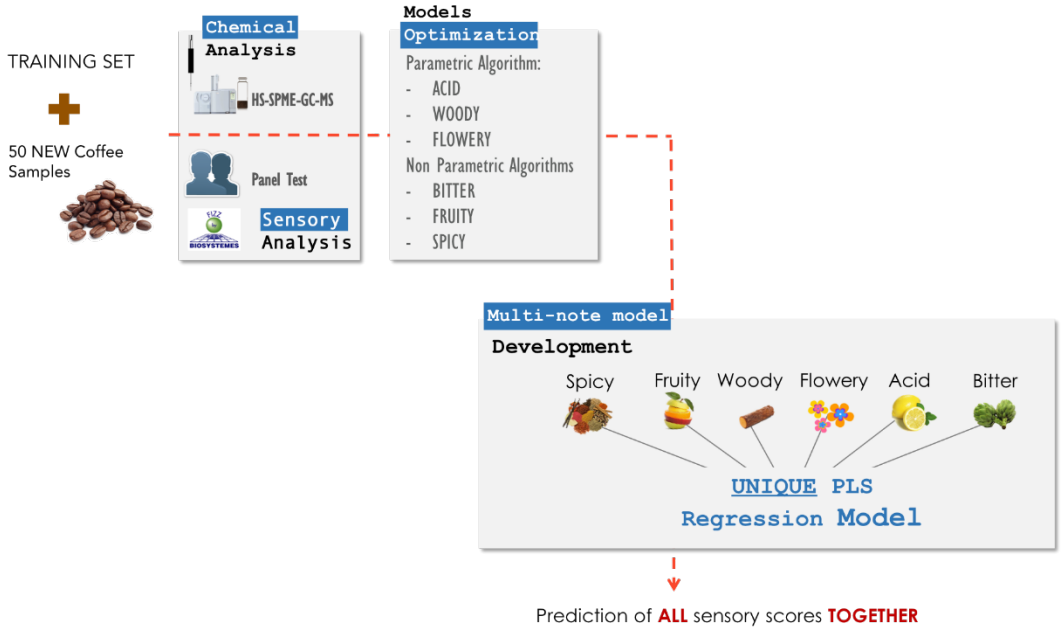
Last but not least, the possibility of automation of the different analytical strategies make the foodomic approach applied and discussed in this thesis suitable as a valid strategy transferable to routine as a complement to the conventional sensory analysis.
The HPLC-UV/DAD analysis implies the preparation of the extracts and their filtration before the injection in the analytical platform while HS-SPME-GC-MS only require the weighting of the samples and then all other steps of the analytical procedure is fully automatic. The possibility of automation makes this second analytical system more suitable to be transferred to routine and quality control.

REFERENCES

1. Spence, C. The multisensory perception of flavour. *Psychologist* **23,** 720–723 (2010).
2. Chambers IV, E. & Koppel, K. Associations of volatile compounds with sensory aroma and flavor: The complex nature of flavor. *Molecules* **18,** 4887–4905 (2013).
3. SCAA. *SCAA Protocols Cupping Specialty Coffee*. *Specialty Coffee Association of America* (2015).
4. Czerny, M., Mayer, F. & Grosch, W. Sensory study on the character impact odorants of roasted Arabica coffee. *J. Agric. Food Chem.* **47,** 695–699 (1999).
5. Grosch, W. Evaluation of the Key Odorants of Foods by Dilution Experiments , Aroma Models and Omission. *Chem. Senses* **26,** 533–545 (2001).
6. Schieberle, P. & Hofmann, T. Mapping the combinatorial code of food flavors by means of molecular sensory science approach. in *Food flavors; chemical, sensory, and technological properties.* (ed. Henryk, J.) 413–438 (CRC Press 2011)
7. Dunkel, A. *et al.* Nature's chemical signatures in human olfaction: A foodborne perspective for future biotechnology. *Angew. Chemie - Int. Ed.* **53,** 7124–7143 (2014).
8. Charve, J., Chen, C., Hegeman, A. D. & Reineccius, G. A. Evaluation of instrumental methods for the untargeted analysis of chemical stimuli of orange juice flavor. *Flavor Fragr. J.* **26,** 429–440 (2011).
9. Abdi, H., Williams, L. J. & Valentin, D. Multiple factor analysis: principal component analysis for multitable and multiblock data sets. *Wiley Interdiscip. Rev. Comput. Stat.* **5,** 149–179 (2013).
10. Blank, I., Sen, A. & Grosch, W. Potent odorants of the roasted powder and brew of Arabica coffee. *Leb. Und-forsch.* **195,** 239–245 (1992).
11. Jeszka-Skowron, M., Sentkowska, A., Pyrzyńska, K. & De Peña, M. P. Chlorogenic acids, caffeine content and antioxidant properties of green coffee extracts: influence of green coffee bean preparation. *Eur. Food Res. Technol.* **242,** 1403–1409 (2016).
12. Martini, S., Conte, A. & Tagliazucchi, D. Phenolic compounds profile and antioxidant properties of six sweet cherry (Prunus avium) cultivars. *Food Res. Int.* **97,** 15–26 (2017).
13. Frank, O., Blumberg, S., Krümpel, G. & Hofmann, T. Structure determination of 3-O-caffeoyl-epi-γ-quinide, an orphan bitter lactone in roasted coffee. *J. Agric. Food Chem.* **56,** 9581–9585 (2008).
14. Frank, O., Zehentbauer, G. & Hofmann, T. Bioresponse-guided decomposition of roast coffee beverage and identification of key bitter taste compounds. *Eur. Food Res. Technol.* **222,** 492–508 (2006).
15. Hofmann, T. Identification of the key bitter compounds in our daily diet is a prerequisite for the understanding of the hTAS2R gene polymorphisms affecting food choice. *Ann. N. Y. Acad. Sci.* **1170,** 116–125 (2009).
16. Flament, I. *Coffee Flavor Chemistry*. (Wiley 2002).
17. Sunarharum, W. B., Williams, D. J. & Smyth, H. E. Complexity of coffee flavor: A compositional and sensory perspective. *Food Res. Int.* **62,** 315–325 (2014).
18. Prescott, J. Multisensory processes in flavor perception and their influence on food choice. *Curr. Opin. Food Sci.* **3,** 47–52 (2015).

2.5 Development of a practical tool to define coffee sensory quality in routine controls

0. GRAPHICAL ABSTRACT



1. INTRODUCTION

Deal with coffee is not as simple as with other agricultural commodities; even when considering the same species and post-harvesting treatment, coffees from different origins present peculiar sensory profiles that need to be qualified to exploit properly their flavour potentialities.
This product unicity lies on the sensory profiles typical of each coffee that are the final results of many factors like species, treatment, origin, seasonality and so on.
Roasters thereby constantly cup and grade the coffees they buy to determine the best way to use them in their blends, or to discover which coffees are of sufficiently high quality to be sold as "single origin" coffees.[1]
Due to its complexity coffee cupping and grading is performed by highly trained panellists able to perceive and discriminate even minor differences among the samples.
Despite its importance, the conventional coffee cupping presents two main drawbacks: it is relatively subjective, and it is time consuming.
The first drawback is related to the high complexity of coffee flavour perception that is related to many factors as previous experiences, memory and unconscious cognitive perceptions and not only to the "raw" sensory stimulation (a better discussion of flavour perception can be found in session 1.1). These limitations can be overcome with a large number of professional panellists that replicate sensory analysis several times to make robust the results. All these requirements support the second drawback of the panel sensory analysis: it is time consuming and not always in compliance with the industrial needs in term of number of samples screened.

The scientific community made a lot of efforts to develop analytical tools to complement and partially replace the human sensory profiling using both separative[2,3] and non-separative[4] techniques coupled to a statistical and multivariate data elaboration. Chemometric data treatment opened new perspectives for coffee quality control affording to consider the overall chemical profile instead of focusing only on well-established targets. This approach found application for coffee quality control not only in the sensory profiling but also in the investigation of coffee adulterations and off-notes[5,6].
Recent studies show a good potential in predicting the sensory scores of coffee samples although their main limit rely on the low number of samples considered and in their variability[2,4,7]. The sensory profiles of coffee is extremely variable; in this perspective, the ambitious aim of this PhD is to characterize chemically, detect and model the variation (in terms of score) of seven different sensory notes within a sample set able representative of this huge variability.

In a previous part of the project (section 2.2) the relationship between chemical composition and sensory impact of different coffee samples has been defined and modelled by a sensometric approach with promising results. The aim of this second step is to strengthen the prediction models and evaluate their flexibility by adding samples covering also different seasonality and new crops and origins and then testing the models previously developed with a pool of new test samples. Since data elaboration is crucial in this respect[8,9] each prediction model previously developed has been optimized and in some cases redefined also using different algorithms that better fit with the data distribution[10].
Last but not least, a multi-note prediction model has been developed where all sensory notes are modelled at the same time, as a panel does at industrial level. Despite the compromises, the big advantage of this single model lies on the possibility to obtain a comprehensive evaluation of the coffee quality in cup with just a single elaboration. This elaboration and its optimization, is important to evaluate if the sensometric approach can be considered complementary to the panel and if these data can be useful to support the conventional sensory analysis.

Although the defined fingerprints and the related prediction models make statistical inference on the test set with an acceptable degree of confidence, these preliminary results need to be strengthened increasing the number of samples or refining the sensory definition of some particular notes.
To develop an instrumental tool able to predict the coffee sensory profiles comparable to those assigned by the panel is complex and challenging; many factors contribute to the complexity as the changes of the coffee over time, the panel alignment and stability, the number of samples that can be analysed with both the techniques. The replacement of the panel contribution with an instrumental data is out of the purposes of this project since the human contribution to the description of a comfort food such as coffee is yet fundamental. However, the results achieved suggest that with a fine tuning from both the sensory and analytical sides, the sensometric approach may be a valid tool for routine screenings.

2. MATERIALS AND METHODS

2.1 Samples Description

A total of 157 coffee samples were collected, roasted, and submitted to chemical and descriptive sensory analyses. More details on samples chemical and sensory analyses are reported in session 2.2 (table 1APX).

The investigated pool of samples includes different species, origins and treatments and was intended to represent a significant although partial variability within the coffees a company manages.
The regression models were built up for six sensory notes (*Acid*, *Bitter*, *Woody*, Flowery, *Fruity* and *Spicy*). The mean values of the scores indicated by the cuppers were used as the dependent variables (y) and the analytical responses of the pool of compounds detected and identified over a set of 364 chromatograms were used as independent variables (x matrix).
The calibration sets consisted of 157 selected samples (314 chromatograms) while the remaining 50samples (100 chromatograms) have been used as an external test set.
The external test set includes not only samples of the same origins of those used in the calibration set but also samples of completely different origins. All samples have been harvested in the year after the training set. This set has been designed to test the models in a “real” practical condition where they are supposed to work independently from origin and harvesting year.
Within the training set 20 randomly-selected samples have been used as a cross-validation set.
The variable selection for the construction of the models was carried out by PLS-DA as previously described in session 2.2.

2.2 Models’ evaluation parameters

The predictive models developed on the training set have been evaluated on the test set considering the following parameters[11]:

- Q^2: a predictive quality index (cross-validated R^2),
- R^2: Determination Coefficient of the Model;
- RMSECV; Root Mean Squared Error in Cross validation
- RMSEP: Root Mean Squared Error in Prediction

All these parameters have been discussed at the end of section 1.3.

3. RESULTS AND DISCUSSION

3.1 Sensory notes modelled through Partial Least Square regression (PLS).

All sensory notes have been modelled through a PLS algorithm; the evolution of each sensory note over the sample sets has been related a different number of variables (Table 1) selected after a PLS-DA elaboration reported in the session 2.2.

Table 1 PLS models' parameters

Sensory note	*n° variables*	Q^2	R^2	*RMSECV*	*RMSEP*
Acid	31	0.680	0.802	0.668	0.720
Bitter	35	0.649	0.656	1.013	0.976
Woody	37	0.750	0.849	0.893	1.260
Flowery	37	0.429	0.444	0.920	0.760
Fruity	23	0.170	0.635	0.900	2.280
Spicy	20	0.178	0.322	1.721	1.314

According to Table 1, *Acid*, *Bitter*, *Woody*, and to a less extend *Flowery* reported PLS prediction models show a good performance. The R^2 values indicate that nearly the 70% of the variance in the measured scores is explained by the models (i.e. the chemical fingerprint used to describe the model) a quite good result in consideration of the high variability of the training set. Moreover, with the exception of the *Flowery* note model, all these sensory notes show high Q^2 values, meaning a good model stability and a good predictive capability. The goodness of the predictive capability is confirmed by the acceptable values of the root mean squared errors (RMSECV and RMSEP) reported for these notes. The limit of acceptability for predicted values has been fixed by the sensory panel, in ± 1 score points. All RMSECV and RMSEP values are within or close to this interval.

Fruity and *Spic*y notes models do not show satisfying performance although some parameters are within the limits of acceptability (i.e. some R^2 values). The main issues are present in the Q^2 values responsible of the predictive quality of the model. This low predictive quality is confirmed by the high values registered for RMSECV and RMSEP.

The low model stability registered for *Flowery, Fruity* and *Spic*y notes might be linked to the unbalanced distribution of the samples within the training set; the number of samples with a low score for these notes is much higher than that with high score. This unbalanced sample distribution makes the use of PLS algorithm challenging in the modelling of these sensory notes requiring more relaxed algorithms able to better follow the evolution of the data matrix.

3.2 Improvement of the critical note fingerprints.

By definition the PLS models finds a linear regression between the predicted and the observed variables by the creation of new variables (similar to PCA) able to explain the variability of the original data matrix. The main PLS limit observed is that it lies on a linear regression not always followed by experimental data.

Non-parametric regression tools afford to summarize a relationship between dependent (sensory scores) and independent variables (Note Related Compounds) with few assumptions about the characteristics of the relationship (e.g. a linear or exponential relationship). In other words, non-parametric regression defines the relationship on the basis of the data distribution in the space and not from a specific trend.

The non-parametric regression algorithm more frequently used is "LOWESS" (LOcally WEighted regression and Smoothing Scatter plots) curve[12]. The acronym gives the idea of a locally weighted regression curve whose values, at a specific location along the x-axis, is determined by the neighbour points. The method thereby does not require assumptions about the form of the relationship and affords to discover the trend with the data itself.
Note fingerprints were elaborated with the LOWESS regression; the modelling of *Flowery*, *Fruity* and *Spicy* significantly improved. Table 2 reports the performances obtained with the LOWESS modelling.

Table 2 Models' performance improved with the Non-parametric algorithm

Sensory note	*n° variables*	Q^2	R^2	*RMSECV*	*RMSEP*
Flowery	37	0.726	0.656	0.881	1.552
Fruity	23	0.678	0.622	0.771	1.541
Spicy	20	0.891	0.746	1.040	1.371

The increasing of R^2 values indicates that the models provide a better explanation of the data matrix variance. These better fitting values are also reflected in the reduction of the Root Mean Squared Error (RMSECV) in the model calibration and, more important, in a reduction of the mean error in the prediction on the test set (RMSEP).
Since the Non-Parametric regression algorithm does not need to fit a pre-existing curve, the homogenous sample distribution between high and low scores is less important in the modelling procedure. This flexibility allowed to improve significantly the model representativity in particular for the *Fruity* and the *Spicy* notes.
The *Flowery* note presents a little different behaviour: the non-Parametric model show better performance in calibration (Q^2, R^2 and RMSECV) but not in prediction; the RMSEP is higher than that calculated with PLS algorithm. This unexpected behaviour may be due to the test set, that, for this particular note, is not perfectly representative of the pool of the training set.

3.3 The *Nutty* Situation

Data collected, and the approach adopted for all other notes was unsuccessful in defining the chemical fingerprint of the *Nutty* note. Several algorithms have been tested to fit the evolution of the scores on the chemical data but none of them gave acceptable results in terms of performance and predictive capability, proving that the *Nutty* note issues are upstream the modelling phase. The list of compounds selected through the PLS-DA were not representative of a high/low expression of *Nutty* note. Since this compound selection strategy was successful for all other notes, the issue may lie on the sensory descriptive analysis of *Nutty*; it is well known that flavour is a cognitive construct[13,14] not only linked to the pure sensation but also on the previous experience of the panellist(s). For *Nutty* this cognitive portion of the perception may have a more important role in the scores assignment compared to the other notes. This aspect would require an important work on the panel training to focus their judgment mainly on the specific *Nutty* description and not on the overall *Nutty* perception
Similar but less accentuated up-stream issues may be at the basis of the high prediction errors and the general worse performance of the *Fruity*, *Flowery* and *Spicy* notes prediction models.

3.4 Observations on the prediction of the external test set

This section reports a critical evaluation of the test-sample sensory scores predicted by the models. Each sensory note has been predicted using the optimized model.

Table 3 Test set prediction summary

	Acid	*Bitter*	*Flowery*	*Fruity*	*Woody*	*Spicy*
Algorithm	PLS	NP-LOWESS	PLS	NP-LOWESS	PLS	NP-LOWESS
% correctly predicted samples	84%	70%	74%	72%	76%	72%

The good performance showed in the previous section have been confirmed also on the fully external test-set. All notes show a mean error in prediction within the acceptability limit and about 70% of the test samples have been successfully predicted. The best model performance has been with the *Acid* note, and the worse predictive capability with the *Fruity* note (Table 3).

Though the overall results are positive, the score points predicted outside the limits of acceptability (± 1) and the reasons of these failures have to be understood and, where possible, the results improved.

Two general considerations can be done for these out of range predicted samples:

- the predicted scores out of the acceptability range were not far from it; all the test set has been predicted within the ±2 score points interval, meaning that the model needs a better accuracy that can be obtained by a more balance and increased number of samples.
- in most inaccurate predictions, the scores of the panel for the test samples were not aligned to those measured for those of the training set. This observation has been done by comparing the scores of the training samples considered as replicates (i.e. belonging to the same origin and/or sub-origin) to those of the test sets. Since, these samples can chemically be considered as replicates (the variability in terms of RSD% is within the instrumental limits) the reason(s) of these discrepancies may be due to: i) the samples are different and the year of harvesting impact massively on the expression of specific sensory notes, ii) the panel need to be re-aligned to reduce their variability with samples that are replicates.
- in some cases, samples used to train the models show a great variability (RSD%= 37 and 47%) of the scores of the panel. This variability compared to an instrumental variability of about 20%, made the note modelling more complex.

A good example to support these observation is provided by Table 4 reporting the samples in which the *Woody* note was not predicted correctly. The same observations can be extended to the other sensory notes (data not shown).

Table 4 Woody critical test samples predictions. Column description: Code = Sample Identification number; Identification: Origin and sub-origin identification code; Meas: Sensory score measured by the panel; Mean Meas: mean of the scores the panel assigned to each origin; Pred: Sensory scores predicted by the models; Err: Error in prediction referred to the single sample measurement (in red when >1). Samples in Blue have been used in the training set while samples in yellow are in the test samples.

Code	Identification	Meas	Mean Meas	Pred	Err
80	JAVA LB	2.80	2.80		
D3_9	JAVA LB	0.00		2.78	2.78
1_VIEGR2_RN	VIETNAM GR 2	5.10			
19	VIETNAM GR 2	4.38			
2_VIEGR2_RN	VIETNAM GR 2	4.28	4.62		
3_VIEGR2_RN	VIETNAM GR 2	3.65			
56	VIETNAM GR 2	6.63			
81	VIETNAM GR 2	3.70			
D3_10	VIETNAM GR 2	1.13		3.94	2.81
D3_50	VIETNAM GR 2	5.20		4.21	0.99
71	UGANDA DRUGAR UTZ ORGANIC	0.00	0.00		
D3_13	UGANDA DRUGAR UTZ ORGANIC	0.00		1.87	1.87
D3_24	UGANDA DRUGAR UTZ ORGANIC	1.87		2.18	0.32
1_BUK_RN	BUKOBA	3.03			
103	BUKOBA	5.23			
116	BUKOBA	3.27			
2_BUK_RN	BUKOBA	6.25			
20	BUKOBA	4.68			
21	BUKOBA	3.40	4.10		
3_BUK_RN	BUKOBA	4.50			
32	BUKOBA	5.28			
48	BUKOBA	2.48			
89	BUKOBA	1.40			
94	BUKOBA	5.64			
D3_25	BUKOBA	2.67		3.85	1.19
D3_46	BUKOBA	1.85		3.55	1.7
D3_14	BUKOBA	2.73		4.58	1.86
D3_41	BUKOBA	3.27		2.88	0.39
17	VIETNAM 18WP	1.63	1.28		
46	VIETNAM 18WP	0.93			
D3_19	VIETNAM 18WP	1.62		3.79	2.17
D3_44	VIETNAM 18WP	3.75		4.98	1.23
1_UGA_RN	UGANDA 18 UP	7.18			
2_UGA_RN	UGANDA 18 UP	6.05			
22	UGANDA 18 UP	3.00			
3_UGA_RN	UGANDA 18 UP	4.20	4.16		
49	UGANDA 18 UP	1.90			
87	UGANDA 18 UP	2.25			
99	UGANDA 18 UP	4.58			
D3_20	UGANDA 18 UP	1.12		4.50	3.38
73	MESSICO ROBUSTA	1.95	1.95		
D3_30	MESSICO ROBUSTA	3.70		3.51	0.19
D3_40	MESSICO ROBUSTA	1.50		4.45	2.95

3.5 Development of single multi-note prediction model

In the development of a useful tool for coffee sensory evaluation the practical aspects of data elaboration are not negligible. Despite the promising results obtained on the test set for modelling the single sensory attribute, the prediction of a coffee complete sensory profile still requires six

different data elaborations. This important data processing is not well in compliance with the routine application of the sensometric approach.
In order to overcome this limitation a single multi-note sensory scores prediction model, able to make statistical inference on all the sensory notes at the same time, was developed. This model would make possible a drastic reduction of the data handling time because it works with a single data matrix in which all sensory scores are included. The variable selection for this multi-note model has been performed by combining the matrices (chemical fingerprints) used for the single-note prediction models. After a first step, when all chemicals (without any repetitions when present on multiple note chemical fingerprint) have been kept, the x matrix has then been simplified and the number of variables reduced according Variable Impact of Projection (VIP) values. This variable reduction has been carried out to reduce the statistical noise and maximise the information provided by each single note chemical fingerprints.
Based to the VIP values (VIP>0.93), 35 compounds on 56 were retained in the building up of the model affording to simplify the multidimensional structure of the prediction model with a negligible loss in performance.
In the development of the multi-note prediction model both the algorithms (PLS and Non-parametric regression) have been tested but PLS provided the best results, although with some compromises in function of the sensory attribute to be described. In particular, these compromises were required from the notes that, singularly, required a Non-Parametric regression.

Table 5 Multi-note model performance summary

	Q^2	R^2	*RMSECV*	*RMSEP*
Acid	0.738	0.743	0.765	0.829
Bitter	0.651	0.661	1.008	0.997
Flowery	0.396	0.401	1.180	1.421
Fruity	0.259	0.277	0.991	1.944
Woody	0.651	0.665	1.334	1.128
Spicy	0.308	0.330	1.691	1.359

As expected, the model has good performance for the notes that were singularly well modelled by a PLS algorithm (*Acid*, *Bitter*, *Flowery* and *Woody*) and less good results for those notes requiring a non-parametric approach *(Fruity* and *Spicy*) (Table 5).
The inclusion of all sensory notes in the same data matrix (and the chemical information related to the pool of compounds needed for their modelling) entailed an increase of the noise and thereby a poorer performance if compared to the single-note models.
The performance in calibration and in prediction (RMSECV and RMSEP) of the multi-note model are not far from those observed in the single prediction models.

4. CONCLUSIONS

The main objective of this part of the project was to test the reliability of the sensory notes chemical fingerprints defined by a sensometric approach in a real application.
The sensory scores of a totally new pool of samples have been predicted using the chemical fingerprint defined in the previous parts of the project (session 2.2.1), also by applying non-parametric algorithm other than PLS in the data modelling.
Each single-note optimized model has been tested on the new samples with encouraging results:

- the models show acceptable performances both in calibration and in prediction: most of the models explain a good portion of the variability around the dependent variable (y) and the Q^2 values demonstrate a good stability of the models. Root Mean Squared Errors in calibration (RMSECV) and in prediction (RMSEP) are below or not far from the limit of acceptability (fixed with the panel in ± 1 score points).
- The main issues in prediction, highlighted with the test set, derived from a rather low stability of the sensory evaluations of the new samples compared to those used to build the models.
  This discrepancy might be due to the impact of the seasonality on the coffee sensory profile. This variability can be better represented in the model development by including in the training phase more coffee samples belonging to the same origin and sub-origin harvested in different years. This step would be time consuming but will give an idea on how and how much the coffee sensory profile changes over the years and according to the climatic changes.

The last part of the project has been focused in the development of a single regression model able to predict the sensory scores of all the notes together one single data elaboration (multi-note model). This model merges and summarizes the whole information extracted by the training samples and points out the compounds globally responsible for the sensory profile descripted by *Acid*, *Bitter*, *Flowery*, *Fruity*, *Woody* and *Spicy* sensory notes. The use of one single model to predict the scores of all these six sensory notes enables to predict the global coffee sensory profile but, it requires some compromises in terms of model robustness and errors in prediction in the case of *Flowery*, *Fruity* and *Spicy* notes while *Acid*, *Bitter* and *Woody* notes are in-line with the performances of the single note models. The evolution of these sensory scores in the training set do not follow a linear behaviour and the best performance were therefore with non-parametric algorithms. Nevertheless, the performance registered in the multi-note prediction model remains acceptable also for *Flowery, Fruity* and *Spicy* notes because the sample distribution for these notes in unbalanced towards the low scores. On the other hand, the linear PLS model overcomes these limits and well represents them.
In addition to these promising results, in view of a routine implementation of this sensometric strategy as a tool complementary to the panel, some limitations have to be overcome with further work:

- coffee, such as many others natural products, presents an extremely high variability; therefore, the number of samples considered in this project is not yet sufficient to obtain reliable and robust models. The training set needs to be extended including more replicates of origins and sub-origins normally treated by a coffee industry and samples harvested during several years (i.e. 5).
- The highest errors in prediction occurred for samples with high sensory scores; a first possible explanation of this behaviour concerns the mathematical model development: the number of high scores samples is lower than that of low scores samples, therefore, this part of the score range is less represented in the sample sets. The second explanation is

related to the stability of the sensory scores measured by the panel: high scores are more difficult to be defined then the low ones and are affected by a stronger variability. The extension of the samples set will provide a good coverage of the whole score range thus overcoming these issues.

- The sensory definition of some notes (i.e. *Nutty*) is challenging. To have stable and reliable scores suitable to be compared to the chemical data, a specific note panel alignment is needed.

Although all these considerations and limits, we can conclude that the Sensometric approach demonstrated to be very promising to characterize coffee sensory notes and can successfully be used to develop tools complementary to the conventional panel analyses, and, after further development, adopted for quality control purposes.

REFERENCES

1. ritta Folmer. The Craft and Science of Coffee. in *The Craft and Science of Coffee* (eds. Blank, I., Farah, A., Giuliano, P., Sanders, D. & Wille, C.) (Elsevier, 2017)
2. Ribeiro, J. S., Augusto, F., Salva, T. J. G., Thomaziello, R. A. & Ferreira, M. M. C. Prediction of sensory properties of Brazilian Arabica roasted coffees by headspace solid phase microextraction-gas chromatography and partial least squares. *Anal. Chim. Acta* **634,** 172–179 (2009).
3. Ribeiro, J. S., Augusto, F., Salva, T. J. G. & Ferreira, M. M. C. Prediction models for Arabica coffee beverage quality based on aroma analyses and chemometrics. *Talanta* **101,** 253–260 (2012).
4. Lindinger, C. *et al.* When Machine Tastes Coffee: Instrumental Approach To Predict the Sensory Profile of Espresso Coffee. *Anal. Chem.* **80,** 1574–1581 (2008).
5. Toledo, B. R., Hantao, L. W., Ho, T. D., Augusto, F. & Anderson, J. L. A chemometric approach toward the detection and quantification of coffee adulteration by solid-phase microextraction using polymeric ionic liquid sorbent coatings. *J. Chromatogr. A* **1346,** 1–7 (2014).
6. Lindinger, C. *et al.* Identification of Ethyl Formate as a Quality Marker of the Fermented Off-note in Coffee by a Nontargeted Chemometric Approach. *J. Agric. Food Chem.* **57,** 9972–9978 (2009).
7. Ongo, E. *et al.* Chemometric discrimination of philippine civet coffee using electronic nose and gas chromatography mass spectrometry. *Procedia Eng.* **47,** 977–980 (2012).
8. Yi, L. *et al.* Chemometric methods in data processing of mass spectrometry-based metabolomics: A review. *Anal. Chim. Acta* **914,** 17–34 (2016).
9. Boccard, J. & Rudaz, S. Harnessing the complexity of metabolomic data with chemometrics. *J. Chemom.* **28,** 1–9 (2014).
10. Martínez-Arranz, I. *et al.* Enhancing metabolomics research through data mining. *J. Proteomics* **127,** 275–288 (2015).
11. Abdi, H. Partial least squares regression and projection on latent structure regression (PLS Regression). *Wiley Interdiscip. Rev. Comput. Stat.* **2,** 97–106 (2010).
12. Cleveland, W. S. Locally Weighted Regression and Smoothing Scatterplots. *J. Am. Stat. Assoc.* **74,** 829–836 (1979).
13. Secundo, L., Snitz, K. & Sobel, N. The perceptual logic of smell. *Curr. Opin. Neurobiol.* **25,** 107–115 (2014).
14. Prescott, J. Multimodal Chemosensory Interactions and Perception of Flavor. in *The Neural Bases of Multisensory Processes* (eds. Murray, M. & Wallace, M.) (2012).

3. Conclusions

Foodomics is an innovative approach capable to characterize food in holistic way and relate the food chemical composition to its biological impact through advanced analytical platforms and multivariate data handling tools. The demand of knowledge and control of the biological impact of what we eat and drink is growing at the same time of consumer awareness of the importance of food composition.
The sensory impact is increasingly important in the product choice, in particular, it is crucial for those foods without a nutritional value but mostly consumed because of the pleasure they elicit known as “comfort foods”. The market of several comfort foods and in particular of coffee is strongly conditioned by the sensory aspects of the products often driven by marketing. The sensory properties of a finish-end coffee products mostly depend on the overall supply chain and, for coffee brands, on the sensory quality of the incoming raw beans, and on their selection in order to characterize and keep a “specialty coffees” (i.e. coffee without defects) to be used as mono-origin or to be mixed to obtain a desired blend with a peculiar flavour.
This research project aimed to define a relationship between sensory and chemical profiles into a coffee sensory evaluation scenario with the final goal to exploit this relationship as a practical tool for coffee sensory quality characterization and to use it as a complement or, when possible, as an alternative to the panel in routine controls. In this context, the international protocol to taste coffee in the selection of the raw beans (SCAA cupping), also known as Brazilian cupping, has to be taken into account as the standardized method in coffee evaluation. The analytical procedure had to be in compliance with this kind of investigation:

- The official SCAA cupping protocol has been simulated by three different analytical approaches (HS-SPME-GC-MS of the roasted powder, HS-SPME-GC-MS of the brew and SBSE of the brew). Chemical data have been compared through chemometric tools and HS-SPME-GC-MS of the powder has been chosen as platform of election for the samples aroma characterization.
  This platform indeed satisfied the requirements in terms of characterization power, easiness and possibility of automation needed for the purposes of the project.

- A Sensometric approach has been used to link the peculiar sensory profile (measured by a trained panel) of different coffee samples to the peculiar chemical composition accurately defined by HS-SPME-GC-MS.
  The sensometric approach proved to be discriminative, informative, and predictive in revealing the chemical signature of the different coffee aroma notes.
  The developed regression models show promising results when applied to predict the samples sensory scores from chemical data validating the chemical signature defined for most notes.

- The reliability of the sensometric approach has been validated by comparing the note characterization obtained with this approach to those obtained by molecular sensory science taken as reference. Despite the differences between the two strategies a good consistency in samples characterization has been found.

- The possibility of improving the knowledge, and thereby the predictive capability of the models when applied to sensory notes known as “taste attributes” has been tested by integrating the HS-SPME-GC-MS chemical data (characterizing the Aroma) with those from a HPLC-UV/DAD fingerprinting analysis targeted on those compounds responsible of these taste-related attributes (e.g *Bitter*).

This integration did not bring to significant improvement of the predictive models developed considering only the volatile fraction, thus confirming the key role of aroma compounds in flavour modulations. On the other hand, the considered taste data make the whole protocol more complex and not suitable for routine implementation.

- The predictive models have been optimized and used to predict the sensory profiles of a completely new set of samples with promising results. All these information have been combined into a single multi-note prediction model that, although with some limits, has shown to be a valid supporting tool for the panel if optimized by integration with a suitable number and kind of samples associated to a fine tuning of the panel.

Considering all these results and the experience acquired over these three years of PhD some general considerations can be done.

Coffee is an extremely fascinating and complex matrix; the chemical characterization of its sensory profiles is still challenging because of the wide variability of this product.

The difficulty of this project is mainly related to the magnitude of the variance that have to be managed by whoever have to deal with this food matrix though we attempted to standardize many of critical parameters (only mono origins were considered; the roasting degree was fixed; the brewing method etc…)

On the other hand, a too rigid management of this variability brings the study out of the real coffee world and risks to make it a pure analytical exercise.

Over the whole project, the samples selection has been carried out with the aim to balance the requests of a quality control laboratory and the need to be representative of such a complex product.

Although our efforts, the number of samples considered is not sufficient to provide a robust and reliable tool for coffee evaluation. The shortage of the number of samples is in particular evident for high scores samples that probably are a minor portion but are equally (if not more) important to model statistically a sensory note.

In addition, the impact of the harvesting period and of the climate differences on the coffee sensory profiles has been considered only in the final part of the project. In the perspective of developing a tool suitable for routine coffee sensory screening, these parameters have to be included in the training set to build the models.

A further consideration arises from the need of close cooperation with the panel: not all sensory notes can be defined with the same degree of confidence.

The unambiguous definition of *Nutty* and *Flowery* note is particularly challenging because of the different facets of these perceptions. The development of the predictive models for these notes is thereby particularly challenging, and modeling of attributes such as body, astringency, aroma intensity etc…. must be considered. The improving of analytical results needs also to be supported by refining sensory analysis in terms of lexicon and stability.

The broad and comprehensive sample characterization achieved by the foodomic approach may open new perspectives in the investigation of the biological impact (even not linked to the senses) of food flavour compounds.

SCIENTIFIC PUBLICATIONS

- **Bressanello, D**., Liberto, E., Collino, M., Reichenbach, S. E., Benetti, E., Chiazza, F., … Cordero, C. (2014). "Urinary metabolic fingerprinting of mice with diet-induced metabolic derangements by parallel dual secondary column-dual detection two-dimensional comprehensive gas chromatography" Journal of Chromatography A, 1361, 265–276.

- Nicolotti, L., Cordero, C., **Bressanello, D**., Cagliero, C., Liberto, E., Magagna, F., … Bicchi, C. (2014). "Parallel dual secondary column-dual detection: A further way of enhancing the informative potential of two-dimensional comprehensive gas chromatography" Journal of Chromatography A, 1360, 264–274.

- Griglione, A., Liberto, E., Cordero, C., **Bressanello,** D., Cagliero, C., Rubiolo, P., … Sgorbini, B. (2015). "High-quality Italian rice cultivars: Chemical indices of ageing and aroma quality". Food Chemistry, 172, 305–313

- Reichenbach, S. E., Rempe, D. W., Tao, Q., **Bressanello, D**., Liberto, E., Bicchi, C., … Cordero, C. (2015). Alignment for Comprehensive Two-Dimensional Gas Chromatography with Dual Secondary Columns and Detectors. Analytical Chemistry, 87(19), 10056–10063

- **Bressanello, D**., Liberto, E., Cordero, C., Rubiolo, P., Pellegrino, G., Ruosi, M. R., & Bicchi, C. (2017). "Coffee aroma: Chemometric comparison of the chemical information provided by three different samplings combined with GC – MS to describe the sensory properties in cup." Food Chemistry, 214, 218–226.

- Belhassen, E., **Bressanello D**., Merle, P., Raynaud, E., Bicchi, C., Cordero, C.) “Routine quantification of 54 allergens in fragrances using comprehensive two-dimensional gas chromatography- quadrupole mass spectrometry with dual parallel secondary columns. Part I : Method development” Flavour and Fragrance Journal, 63–74 (2017)

- **Bressanello D**, Liberto E., Collino M, Chiazza F, Mastrocola R., Reichenbach S.E, Bicchi C. and Cordero C. “Combined Untargeted and Targeted Fingerprinting by Comprehensive Two-Dimensional Gas Chromatography: Revealing Fructose-Induced Changes in Mice Urinary Metabolic Signatures” Analytical and Bioanalytical Chemistry IN PRESS

- **Bressanello D**, Liberto E, Cordero C, Sgorbini B., Rubiolo P., Pellegrino G., Ruosi M.R., Bicchi C. “Chemical signature of coffee aroma notes: an objective tool for quality control” IN PREPARATION FOR THE SUBMISSION TO JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

CONGRESSES CONTRIBUTIONS

38th International Symposium on Capillary Chromatography and 11th GCxGC Symposium

- "Exploring the amyloid beta (abeta) dependent metabolic fingerprint of hippocampal neurons by two-dimensional comprehensive gas chromatography: potentials in drug discovery studies"
Davide Bressanello; Davide Garella; Massimo Bertinaria; Andrea Marcantoni; Claudio Franchino; Gianluca Miglio; Stephen E. Reichenbach; Carlo Bicchi; Chiara Cordero
Poster)

- "Tea (camellia sinensis l.) volatiles profiling by headspace – two-dimensional comprehensive gas chromatography-mass spectrometry: flavor compounds quantification challenges"
Federico Magagna; Erica Liberto; Davide Bressanello; Cecilia Cagliero; Patrizia Rubiolo; Barbara Sgorbini; Carlo Bicchi; Sabrina Seno; Marco Franci; Chiara Cordero
(Poster)

- "Dual second dimension column-dual detection in two-dimensional comprehensive gas chromatography (GC×2GC-MS/FID): increased information in optimized separation conditions in metabolomics"
Davide Bressanello; Erica Liberto; Massimo Collino; Stephen E. Reichenbach; Elisa Benetti;Fausto Chiazza; Carlo Bicchi; Chiara Cordero
(Poster)

X Italian Congress of Food Chemistry

- "Virgin olive oil blueprint? As the analytical platform could enhance the fingerprinting sensitivities and effectiveness of sample classification"
Erica Liberto; Chiara Cordero; Davide Bressanello; Carlo Bicchi; Giorgia Purcaro; Lanfranco Conte
(Poster)

- Effect of Storage Time on the Aroma of Italian High Quality Oryza sativa L. Cultivars: Chemical Indices of Ageing
Erica Liberto; Alessandra Griglione; Chiara Cordero; Davide Bressanello; Diego Greppi; Carlo Bicchi
(Oral Presentation)

CoCoTea International Congress 2015

- "Packaging and coffee aroma: a kinetic evolution"
Davide, Bressanello; Erica, Liberto; Chiara Emilia, Cordero; Carlo, Bicchi
(Poster)

40th International Symposium on Capillary Chromatography and 13th GCxGC Symposium riva 2015

- "Coffee aroma investigation: a chemometric comparison of the information provided by three HCC sampling techniques able to chemically describe the sensory properties in cup."
Davide, Bressanello; Erica, Liberto; Chiara, Cordero; Patrizia, Rubiolo; Gloria, Pellegrino; Rosanna, Ruosi manuela; Carlo, Bicchi
(Poster)

- "Human urinary metabolic signatures by parallel dual secondary column-dual detection two-dimensional comprehensive gas chromatography: reliable targeted and untargeted comparative analysis"
Davide Bressanello; Erica Liberto; D. W. Rempe; Q. Tao; Stephen E. Reichenbach; Stefano Balducci; Elisa Benetti; Carlo Bicchi; Chiara E. Cordero
(Poster)

- "Potentials of the reversed-inject differential flow modulator comprehensive two-dimensional gas chromatography in the quantitative profiling of suspected fragrance allergens"
Cordero, Chiara; Bressanello, Davide; Rubiolo, Patrizia; Cobelli, Luigi; Miliazza, Armando; Giardina, Matthew; Bicchi, Carlo
(Poster)

XI Italian Congress of Food Chemistry 2016

- "Sensometric for the chemical odor code characterization of different coffee aroma notes."
Erica, Liberto; Davide, Bressanello; Chiara, Cordero; Barbara, Sgorbini; Cecilia, Cagliero; Patrizia, Rubiolo; Gloria, Pellegrino; Rosanna, Ruosi Manuela; Carlo, Bicchi
(Oral Presentation)

SCAA Expo 2017

- "Chemical signature characterization of different coffee aroma notes."
Bressanello, D.; Liberto, E.; Bicchi, C.; Ruosi, M.R.; Pellegrino, G.
(Poster)

CoCoTea International Congress 2017

- "Coffee aroma sensory notes and chemical signature: a sensometric approach"
Bressanello, D.; Liberto, E.; Bicchi, C.; Ruosi, M.R.; Pellegrino, G.
(Oral Presentation)

- "Coffee 'Identitation' through chromatographic fingerprint: simultaneous classification of geographical origin & post-harvest treatments"
Bressanello, Davide; Tarighat, Maryam Abbasi; Liberto, Erica; Pellegrino, Gloria; Ruosi, Manuela R.; Bicchi, Carlo
(Poster)

- “Flavoromics approach to describe the sensory properties of a cup of coffee”
Bressanello, Davide; Liberto, Erica; Sgorbini, Barbara; Cordero, Chiara; Rubiolo, Patrizia; Pellegrino, Gloria; Ruosi, Manuela R.; Bicchi, Carlo
(Poster)

- “Comparison between sensometric and sensomic approaches in the sensory-chemistry relationship definition”
Bressanello, Davide; Liberto, Erica; Cordero, Chiara; Sgorbini, Barbara; Rubiolo, Cecilia Cagliero Patrizia; Pellegrino, Gloria; Ruosi Manuela R.; Bicchi, Carlo
(Poster)

4. Appendix

Supplementary material relative to chapter 2.2.

Table 1APX Coffee samples and sensory characteristics. Spec: Species (A: Arabica, R: Robusta), Treat: Treatments (N: Natural, W: Washed).

#	Sample acronym	Type	Spec	Trea	Acid	Bitter	Aroma intensity	Flowery	Fruity	Woody	Nutty	Spicy	Body	Astringency
							SAMPLES WITH PECULIAR SENSORY ATTRIBUTES							
1	*BRA*	BRAZIL LA2	A	N	2	2	7	1	0	1	3	0	7	1
2	*COL*	COLOMBIA CL1	A	W	4	1	8	3	3	0	0	0	7	1
3	*JAV*	JAVA WB1 MB	R	W	0	3	8	0	0	3	3	1	8	1
4	*UGA*	UGANDA STD	R	N	0	3	8	0	0	4	3	3	8	1
5	*PNG*	PAPUA NG Y	A	W	3	2	7	3	3	0	0	0	8	0
6	*INDIA*	INDIA ARAB CHERRY	A	N	2	4	8	0	0	2	2	3	8	1
7	*INDO*	INDONESIA EK1	R	N	0	4	8	0	0	5	3	3	8	2
8	*KAFA*	ETIOPIA KAFA GR. 3	A	N	4	1	8	3	6	0	0	0	7	0
							SELECTED SENSORY STRESSED SAMPLES							
9	*BRALA2*	BRAZIL LA2	A	N	4	1	6	0	0	0	5	0	-	-
10	*BRAGOU*	BRAZIL GOURMET	A	N	2	0	5	1	0	0	5	0	-	-
11	*JAV*	JAVA MB	R	W	1	2	7	0	0	1	7	0	-	-
12	*D2_65*	INDO_CN	R	N	0	5	7	0	0	6	0	2	-	-
13	*D2_37*	BRASILE RFA	A	N	3	0	8	0	0	0	8	0	-	-
14	*D2_56*	Vietnam GR 2	R	N	0	4	8	0	0	7	1	1	-	-
15	*INDOEK1*	INDONESIA EK1	R	N	0	4	8	0	0	7	1	3	-	-
16	*INDIACHAB*	INDIA CHY AB	R	N	0	3	7	0	0	7	3	6	-	-
17	*BUK*	BUKOBA	R	N	0	5	8	0	0	5	2	8	-	-
18	*CON*	CONILON	R	N	0	4	8	0	0	4	2	4	-	-
19	*VIEGR2*	VIETNAM GR2	R	N	0	3	6	0	0	4	3	2	-	-
20	*UGA*	UGANDA 18 UP	R	N	0	4	8	0	0	6	1	7	-	-
21	*D2_1*	BURUNDI	A	W	6	0	7	6	1	0	0	0	-	-

#	Sample acronym	Type	Spec	Trea	Acid	Bitter	Aroma intensity	Flowery	Fruity	Woody	Nutty	Spicy	Body	Astringency
22	*D2_2*	COLOMBIA CL1	A	W	4	0	6	3	0	0	0	0	-	-
23	*D2_3*	COLOMBIA CL4	A	W	4	0	7	3	2	0	0	0	-	-
24	*D2_4*	TANZANIA KILIMANJARO	A	W	5	0	7	3	1	0	0	0	-	-
25	*D2_5*	PAPUA NUOVA GUINEA Y GR.1	A	W	4	0	7	2	3	0	0	0	-	-
26	*D2_6*	RWANDA	A	W	5	0	8	3	2	0	1	0	-	-
27	*D2_7*	ETHIOPIA SIDAMO GR.2	A	W	5	0	8	5	4	0	0	0	-	-
28	*D2_8*	TANZANIA AB	A	W	5	0	7	3	3	0	0	0	-	-
29	*D2_9*	BRASILE LS2	A	N	2	0	7	0	0	0	5	0	-	-
30	*D2_10*	BRASILE LAMBARI	A	N	1	1	5	0	0	0	2	0	-	-
31	*D2_11*	BRASILE LB1	A	N	2	0	7	0	0	0	4	0	-	-
32	*D2_12*	BRASILE APPASSITA	A	N	3	0	8	0	0	0	3	0	-	-
33	*D2_13*	BRASILE LA1	A	N	2	0	8	0	0	0	5	0	-	-
34	*D2_14*	BRASILE LA3	A	N	0	2	6	0	0	1	2	1	-	-
35	*D2_15*	INDIA ARABICA	A	N	1	2	4	0	1	1	0	2	-	-
36	*D2_16*	GUATEMALA ROBUSTA	R	W	1	2	7	0	0	1	3	0	-	-
37	*D2_17*	VIETNAM POLISHED	R	W	0	2	8	0	0	2	5	0	-	-
38	*D2_18*	KAAPIROIALE	R	W	1	2	8	0	0	1	5	0	-	-
39	*D2_19*	VIETNAM GR 2	R	N	0	4	6	0	0	4	0	1	-	-
40	*D2_20*	TANZANIA	R	N	0	3	6	0	0	5	0	2	-	-
41	*D2_21*	BUKOBA	R	N	0	3	6	0	0	3	0	3	-	-
42	*D2_22*	UGANDA STD	R	N	0	3	7	0	0	3	2	2	-	-
43	*D2_23*	CONILON	R	N	0	2	7	0	0	2	1	1	-	-
44	*D2_24*	INDONESIA EK1	R	N	0	4	7	0	0	4	0	3	-	-
45	*D2_25*	ETHIOPIA SIDAMO GR.2	A	W	2	0	5	2	0	0	0	0	-	-
46	*D2_26*	PAPUA NUOVA GUINEA Y GR.1	A	W	3	0	6	2	2	0	0	0	-	-
47	*D2_27*	COLOMBIA CL1	A	W	1	0	4	1	0	0	0	0	-	-
48	*D2_28*	BRASILE LS2	A	N	1	0	7	0	0	0	7	0	-	-
49	*D2_29*	BRASILE APPASSITA	A	N	2	0	5	1	0	0	2	0	-	-
50	*D2_30*	GUATEMALA ROBUSTA	R	W	0	1	8	0	0	2	5	0	-	-
51	*D2_31*	KAAPIROIALE	R	W	1	1	8	0	0	1	7	1	-	-
52	*D2_32*	TANZANIA	R	N	0	4	6	0	0	5	0	2	-	-
53	*D2_33*	CONILON	R	N	0	3	7	0	0	4	1	1	-	-
54	*D2_34*	COLOMBIA CL2	A	W	4	0	7	3	2	0	0	0	-	-

#	Sample acronym	Type	Spec	Trea	Acid	Bitter	Aroma intensity	Flowery	Fruity	Woody	Nutty	Spicy	Body	Astringency
55	*D2_35*	COLOMBIA CL3	A	W	5	0	9	4	4	0	0	0	-	-
56	*D2_36*	GUATEMALA HB	A	W	2	3	4	0	0	1	1	0	-	-
57	*VIEGR1*	VIETNAM GR1 16 WP	R	T	1	3	7	0	0	4	4	1	-	-
58	*D2_38*	BRASILE LA5	A	N	1	1	7	0	0	0	4	0	-	-
59	*D2_39*	INDIA ROBUSTA CHERRY AA	R	N	0	3	7	0	0	3	0	2	-	-
60	*D2_40*	VIETNAM GR1 CLEAN	R	C	0	3	7	0	0	3	1	1	-	-
61	*D2_41*	VIETNAM GR1 2%	R	N	0	3	7	0	0	3	1	0	-	-
62	*D2_42*	BRASILE LB1	A	N	2	0	7	0	0	0	3	0	-	-
63	*D2_43*	BRASILE LA1	A	N	3	0	6	0	0	0	2	0	-	-
64	*D2_44*	BRASILE LAMBARI	A	N	1	1	5	0	0	0	1	0	-	-
65	*D2_45*	BRASILE LA3	A	N	1	1	6	0	0	0	2	0	-	-
66	*D2_46*	VIETNAM POLISHED	R	W	0	2	8	0	0	1	6	0	-	-
67	*D2_47*	INDONESIA EK1	R	N	0	4	6	0	0	4	1	1	-	-
68	*D2_48*	BUKOBA	R	N	0	2	7	0	0	2	1	4	-	-
69	*D2_49*	UGANDA STD	R	N	0	2	8	0	0	2	1	7	-	-
70	*D2_50*	COLOMBIA CL2	A	W	3	0	6	3	0	0	0	0	-	-
71	*D2_51*	COLOMBIA CL3	A	W	4	0	6	3	2	0	0	0	-	-
72	*D2_52*	BRASILE RFA	A	N	2	0	7	0	0	0	6	0	-	-
73	*D2_53*	BRASILE LA5	A	N	1	1	6	0	0	1	2	0	-	-
74	*D2_54*	VIETNAM GR1 CLEAN	R	C	0	2	8	0	0	4	1	0	-	-
75	*D2_55*	VIETNAM GR1 2%	R	N	0	3	8	0	0	5	1	0	-	-
14	*INDOAP*	INDONESIA AP	R	T	0	3	7	0	0	5	3	1	-	-
77	*D2_57*	INDIA ROBUSTA CHERRY AA	R	N	0	6	3	0	0	5	0	1	-	-
78	*D2_58*	KENYA_CN	A	L	3	1	6	1	1	0	1	0	-	-
79	*D2_59*	ETHIOP_FR	A	N	1	2	5	0	0	0	5	0	-	-
80	*D2_60*	COL_EXC	A	L	4	0	6	2	0	0	0	0	-	-
81	*D2_61*	PERU_HB	A	L	2	1	4	0	0	0	1	0	-	-
82	*D2_62*	BURUND_CN	A	L	2	1	4	0	0	0	0	0	-	-
83	*D2_63*	BRASIL_GR	A	N	1	0	1	0	0	0	0	0	-	-
84	*D2_64*	CHERRY	R	N	0	3	7	0	0	3	1	5	-	-
85	*KAAP*	KAAPIROYALE	R	W	1	3	8	0	0	3	6	4	-	-
86	*D2_66*	LAOS	A	N	3	1	7	3	1	0	0	0	-	-
87	*D2_67*	LAMBARI	A	N	3	0	8	0	0	0	6	0	-	-

#	Sample acronym	Type	Spec	Trea	Acid	Bitter	Aroma intensity	Flowery	Fruity	Woody	Nutty	Spicy	Body	Astringency
88	*D2_68*	LB3	A	N	2	1	7	0	0	0	4	0	-	-
89	*D2_69*	LA2	A	N	3	1	5	0	0	0	3	0	-	-
90	*D2_70*	LA5	A	N	1	3	2	0	0	0	0	0	-	-
91	*D2_71*	UG_DRUGAR	R	N	2	1	8	3	1	0	0	0	-	-
92	*D2_72*	VIETGR12%	R	N	0	3	7	0	0	4	2	1	-	-
93	*D2_73*	MESSIC_RB	R	N	0	1	8	0	0	2	5	1	-	-
94	*D2_74*	KENYAABC	A	L	2	1	6	1	1	0	2	0	-	-
95	*D2_75*	CL2	A	L	3	0	6	2	1	0	0	0	-	-
96	*D2_76*	CL4	A	L	2	0	7	3	0	0	1	0	-	-
97	*D2_77*	LA1	A	N	2	0	8	0	0	0	7	0	-	-
98	*D2_78*	LA3	A	N	1	1	8	0	0	0	5	0	-	-
99	*D2_79*	CHERRY_AB	R	N	0	3	7	0	0	4	1	2	-	-
100	*D2_80*	JAVA	R	L	1	3	5	0	0	3	2	0	-	-
101	*D2_81*	VIETNAM_GR2	R	N	0	2	8	0	0	4	2	2	-	-
102	*D2_82*	HONDURHB	A	L	2	1	6	1	0	0	0	0	-	-
103	*D2_83*	NICARAG	A	L	1	1	5	0	0	1	0	0	-	-
104	*D2_84*	KAFA	A	N	3	0	7	3	6	0	0	0	-	-
105	*D2_85*	LS2	A	N	2	1	8	0	0	0	6	0	-	-
106	*D2_86*	LA3RFA	A	N	1	2	7	0	0	0	5	0	-	-
107	*D2_87*	UGANDA18	R	N	1	3	8	0	0	2	1	2	-	-
108	*D2_88*	VIETN_3%	R	N	0	3	7	0	0	4	1	1	-	-
110	*D2_90*	CL3	R	N	3	0	6	1	1	0	1	0	-	-
111	*D2_91*	PNG	A	L	2	0	6	3	0	0	0	0	-	-
112	*D2_92*	HONDSHG	A	L	3	1	7	0	1	0	1	0	-	-
113	*D2_93*	KAFA	A	L	3	0	8	2	0	0	0	0	-	-
114	*D2_94*	BUKOBA	A	N	0	3	7	5	1	6	0	0	-	-
115	*D2_95*	INDOEK80	R	N	0	6	5	0	0	7	0	5	-	-
116	*D2_96*	LIMU	A	L	2	0	7	0	0	0	0	2	-	-
117	*D2_97*	KAFA	A	N	3	0	6	5	1	0	0	0	-	-
118	*D2_98*	CHERRYAR	A	N	1	4	5	2	5	5	0	0	-	-
119	*D2_99*	UGANDA18	R	N	0	4	7	0	0	5	0	0	-	-
120	D2_100	KILIMANJ	A	L	1	0	5	0	0	0	0	4	-	-
121	D2_101	KENYA ABC	A	L	1	1	6	0	1	0	3	0	-	-
122	D2_102	INDIA CHERRY ROB AB	R	N	0	4	8	0	0	3	1	6	-	-

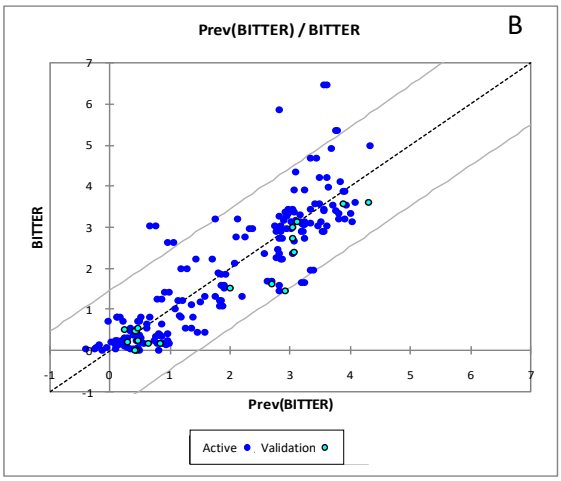
#	Sample acronym	Type	Spec	Trea	Acid	Bitter	Aroma intensity	Flowery	Fruity	Woody	Nutty	Spicy	Body	Astringency
123	D2_103	BUKOBA	R	N	0	4	8	0	0	5	0	5	-	-
124	D2_104	PERU' ORGANICO	A	L	2	0	8	4	0	0	0	0	-	-
125	D2_105	PERU' T1 UTZ/ORG	A	L	3	1	7	0	2	0	2	0	-	-
126	D2_106	INDONESIA EK 1	R	N	0	7	9	0	0	5	0	5	-	-
127	D2_107	INDIA CN	R	N	0	9	9	0	0	8	0	2	-	-
128	D2_108	PERU' TIPO I	A	L	4	0	7	4	1	0	0	0	-	-
129	D2_109	PERU' TIPO II	A	L	3	1	5	3	0	0	0	0	-	-
130	D2_110	MESSICO ARABICA	A	L	2	1	4	0	0	2	0	0	-	-
131	D2_111	INDIA CHERRY ROB AB	R	N	0	4	8	0	0	4	0	4	-	-
132	D2_112	PERU' GOURMET	A	L	2	0	7	3	0	0	0	0	-	-
133	D2_113	GUATEMALA SHB	A	L	3	3	3	5	2	0	0	0	-	-
134	D2_114	TANZANIA AB RFA	A	L	3	0	8	4	6	0	0	0	-	-
135	D2_115	ETHIOPIA YRGACHEFFEE G1	A	L	4	0	9	9	4	0	0	0	-	-
136	D2_116	BUKOBA	R	N	3	4	6	0	0	3	0	3	-	-
137	D2_117	ETHIOPIA YRGACHEFFEE G1	A	L	4	0	7	4	0	0	0	0	-	-
138	D2_118	TANZANIA AB PLUS	A	L	4	0	7	3	3	0	0	0	-	-
139	D2_119	TANZANIA AB PLUS	A	L	2	1	6	1	2	0	3	0	-	-
140	D2_120	COSTA RICA	A	L	4	0	9	1	7	0	0	0	-	-
141	D2_121	COSTA RICA	A	L	4	0	8	3	4	0	0	0	-	-

Bitter

A

Training Set		
Data Matrix	n° Observations	218
	n° of Explicative Variables	21
	n° of Quantitative Dependent Variables	1
Cross Validation	n° Random Chosen Observations	18
Test Set	n° Random Chosen Observations	30

B



C

Observation	Predicted	Measured	Res. (Abs Value)
12_1	0,1847	0,00	0,18
25_2	0,8311	0,00	0,83
35_2	0,1502	0,00	0,15
1_1	0,1890	0,02	0,17
34_1	0,7310	0,10	0,63
7_1	0,7758	0,14	0,64
13_1	0,4285	0,15	0,28
50_2	0,3595	0,20	0,16
5_1	0,3982	0,22	0,18
6_1	0,8001	0,38	0,42
61_1	0,4797	0,53	0,05
85_2	0,3247	0,55	0,23
82_2	0,5555	0,70	0,14
74_1	-0,0595	0,80	0,86
53_1	1,1159	0,83	0,28
45_1	1,1408	1,03	0,12
1_BRALA2_AN_2	1,4304	1,10	0,33
1_JAV_RL_1	1,5304	1,18	0,35
54_2	2,8559	2,47	0,39
20_1	3,9485	3,13	0,82
79_2	3,2248	3,30	0,08
3_KAAP_RL_2	2,8969	3,35	0,45
39_1	2,8489	3,37	0,52
89_2	3,6566	3,55	0,11
1_INDOEK1_RT_1	3,2971	3,98	0,68
3_BUK_RN_1	3,6133	4,10	0,49
1_CON_RN_2	3,1609	4,35	1,19
1_UGA_RN_2	3,5722	4,93	1,36
65_1	4,3869	5,00	0,61
57_2	2,8563	5,87	3,01

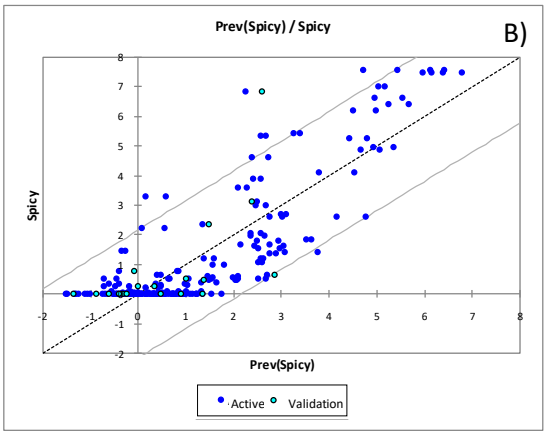
Model Performance Parameters	
R^2	0.721
Q^2	0.689
Median of Error in Prediction	0.443

Spicy

A)

Training Set		
Data Matrix	n° Observations	248
	n° of Explicative Variables	21
	n° of Quantitative Dependent Variables	1
Cross Validation	n° Random Chosen Observations	18
Test Set	n° Random Chosen Observations	30

B)



C)

Observation	Predicted	Measured	Res. (Abs Value)
1_1	-0,6177	0,0	0,618
1_BRALA2_AN_2	0,9622	0,00	0,962
12_1	0,3089	0,00	0,309
13_1	-0,0866	0,00	0,087
25_2	0,4410	0,00	0,441
35_2	-0,1540	0,00	0,154
45_1	-0,2067	0,00	0,207
53_1	-0,0403	0,00	0,040
54_2	0,9207	0,00	0,921
7_1	0,6902	0,00	0,690
74_1	0,1778	0,00	0,178
82_2	0,1692	0,00	0,169
85_2	0,0031	0,00	0,003
5_1	0,3162	0,02	0,296
61_1	-0,0049	0,03	0,030
50_2	-0,3064	0,03	0,340
6_1	-0,2080	0,06	0,268
1_JAV_RL_1	1,3655	0,48	0,885
57_2	1,4573	0,60	0,857
65_1	2,8450	0,63	2,212
39_1	1,5874	1,20	0,387
20_1	3,0802	1,40	1,680
79_2	2,4825	1,80	0,682
3_KAAP_RL_2	2,4224	3,90	1,478
1_INDOEK1_RT_1	2,5823	5,35	2,768
1_CON_RN_2	3,2564	5,45	2,194
89_2	4,5032	6,20	1,697
1_UGA_RN_2	5,6839	6,40	0,716
49_1	2,2674	6,83	4,566
3_BUK_RN_1	6,1234	7,57	1,447

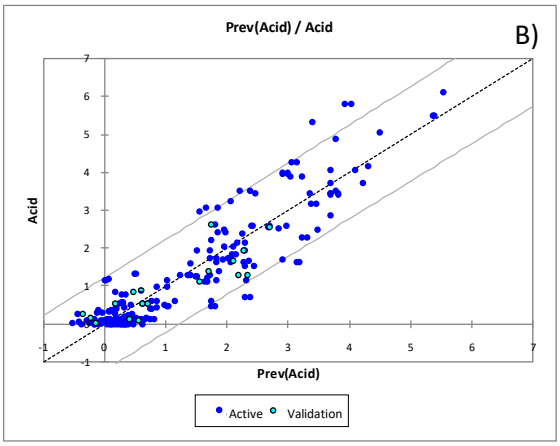
Model Performance Parameters	
R^2	0.684
Q^2	0.231
Median of Error in Prediction	1.072

Acid

A)

Training Set		
Data Matrix	n° Observations	218
	n° of Explicative Variables	24
	n° of Quantitative Dependent Variables	1
Cross Validation	n° Random Chosen Observations	18
Test Set	n° Random Chosen Observations	30

B)



C)

Observation	Predicted	Measured	Res. (Abs Value)
1_CON_RN_2	0,2235	0,00	0,223
1_UGA_RN_2	0,1630	0,00	0,163
20_1	-0,3434	0,00	0,343
57_2	0,1660	0,00	0,166
65_1	0,1506	0,00	0,151
89_2	-0,8055	0,05	0,856
79_2	-0,2024	0,13	0,327
3_BUK_RN_1	0,1906	0,13	0,061
54_2	0,1583	0,13	0,025
39_1	0,5377	0,23	0,304
1_INDOEK1_RT_1	0,4793	0,55	0,071
1_JAV_RL_1	0,9676	1,15	0,182
3_KAAP_RL_2	0,0733	1,18	1,107
53_1	1,6743	1,23	0,441
45_1	1,3182	1,60	0,282
61_1	2,2894	1,63	0,656
25_2	2,9700	1,70	1,270
85_2	1,9167	1,75	0,167
82_2	1,7120	2,20	0,488
13_1	2,5278	2,38	0,153
74_1	3,4876	2,48	1,013
12_1	2,9090	2,53	0,384
50_2	4,0472	2,87	1,181
1_BRALA2_AN_2	1,9502	3,23	1,280
5_1	4,3609	4,16	0,201
35_2	3,8422	4,87	1,024
6_1	4,8677	5,04	0,172
7_1	3,4134	5,30	1,887
1_1	5,3410	6,10	0,759
34_1	3,4881	3,50	0,012

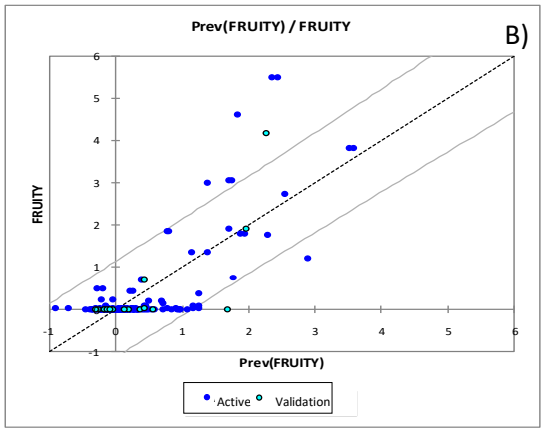
Model Performance Parameters	
R^2	0.831
Q^2	0.497
Median of Error in Prediction	0.543

Fruity

A)

Training Set		
Data Matrix	n° Observations	212
	n° of Explicative Variables	20
	n° of Quantitative Dependent Variables	1
Cross Validation	n° Random Chosen Observations	18
Test Set	n° Random Chosen Observations	30

B)



C)

Observation	Predicted	Measured	Res. (Abs Value)
1_BRALA2_AN_2	-0,1063	0,00	0,11
1_CON_RN_2	-0,3913	0,00	0,39
1_JAV_RL_1	0,1253	0,00	0,13
1_UGA_RN_2	0,2170	0,00	0,22
12_1	1,1367	0,00	1,14
20_1	0,5180	0,00	0,52
3_BUK_RN_1	-0,0058	0,00	0,01
39_1	-0,0294	0,00	0,03
45_1	-0,2517	0,00	0,25
50_2	1,1395	0,00	1,14
54_2	-0,1498	0,00	0,15
57_2	0,1799	0,00	0,18
61_1	-0,0775	0,00	0,08
65_1	0,0691	0,00	0,07
79_2	-0,1510	0,00	0,15
82_2	-0,3667	0,00	0,37
85_2	1,7778	0,00	1,78
89_2	-0,5645	0,00	0,56
1_INDOEK1_RT_1	0,2132	0,03	0,18
53_1	0,1068	0,03	0,07
3_KAAP_RL_2	0,3016	0,08	0,22
13_1	0,3817	0,15	0,23
25_2	0,5335	0,38	0,16
74_1	0,8659	0,75	0,12
1_1	2,5866	1,20	1,39
6_1	0,9475	1,78	0,83
5_1	4,1040	2,73	1,38
34_1	2,3036	3,00	0,70
35_2	2,2529	4,17	1,91
7_1	1,5458	4,63	3,08

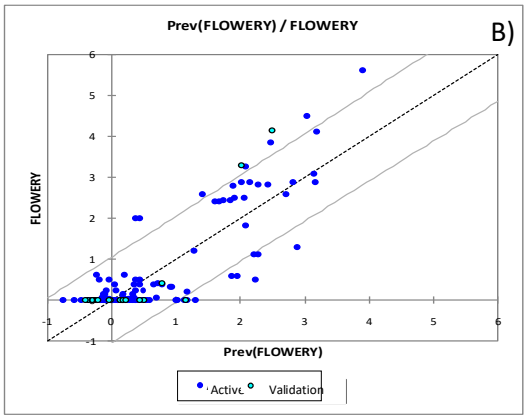
Model Performance Parameters	
R^2	0.661
Q^2	0.193
Median of Error in Prediction	0.206

Flowery

A)

Training Set		
Data Matrix	n° Observations	214
	n° of Explicative Variables	22
	n° of Quantitative Dependent Variables	1
Cross Validation	n° Random Chosen Observations	18
Test Set	n° Random Chosen Observations	30

B)



C)

Observation	Predicted	Measured	Res. (Abs Value)
1_BRALA2_AN_2	-0,1301	0,00	0,13
1_CON_RN_2	-0,6175	0,00	0,62
1_INDOEK1_RT_1	0,2743	0,00	0,27
1_JAV_RL_1	-0,1622	0,00	0,16
1_UGA_RN_2	-0,0091	0,00	0,01
20_1	-0,2901	0,00	0,29
3_BUK_RN_1	0,1107	0,00	0,11
3_KAAP_RL_2	0,3366	0,00	0,34
39_1	0,0683	0,00	0,07
45_1	0,1843	0,00	0,18
53_1	0,3023	0,00	0,30
54_2	0,0356	0,00	0,04
57_2	-0,3807	0,00	0,38
65_1	-0,1526	0,00	0,15
79_2	-0,3728	0,00	0,37
85_2	0,2435	0,00	0,24
89_2	-0,8467	0,00	0,85
12_1	0,5381	0,05	0,49
61_1	1,1522	0,20	0,95
13_1	0,3366	0,38	0,04
74_1	1,8890	0,50	1,39
82_2	1,2664	1,20	0,07
5_1	2,8651	1,30	1,57
26_2	2,2433	1,83	0,42
6_1	2,6053	2,60	0,01
50_2	2,1716	2,80	0,63
8_2	3,1169	3,10	0,02
35_2	2,3828	3,87	1,48
34_1	2,6362	4,50	1,86
1_1	3,8649	5,62	1,76

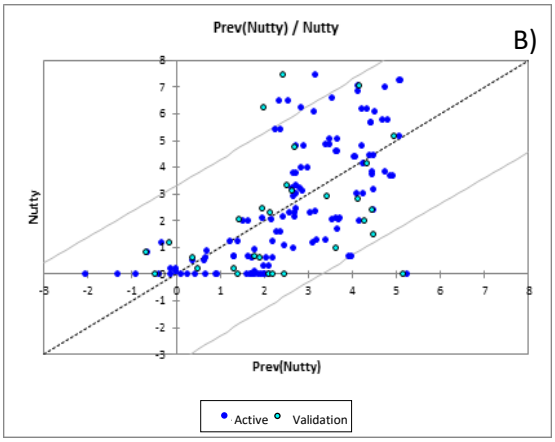
Model Performance Parameters	
R^2	0.793
Q^2	0.342
Median of Error in Prediction	0.351

Nutty

A)

Training Set		
Data Matrix	n Observations	156
	n of Explicative Variables	22
	n of Quantitative Dependent Variables	1
Cross Validation	n Random Chosen Observations	18
Test Set	n Random ChosenObservations	30

B)



C)

Observation	Predicted	Measured	Res. (Abs Value
1_2	-0.67	0.00	0.67
5_1	1.63	0.00	1.63
21_2	1.10	0.00	1.10
25_2	-1.13	0.00	1.13
26_1	1.69	0.00	1.69
32_2	0.70	0.00	0.70
8_1	-0.32	0.02	0.34
20_1	1.11	0.13	0.98
34_2	1.12	0.13	0.99
3_INDOEK1_RT_1	3.85	0.72	3.13
3_INDOEK1_RT_2	4.04	0.72	3.32
56_2	0.86	0.90	0.04
49_1	1.59	0.95	0.64
44_1	1.65	1.10	0.55
3_UGA_RN_2	4.09	1.20	2.89
33_2	1.92	1.25	0.67
1_UGA_RN_2	4.20	1.50	2.70
2_UGA_RN_2	3.85	1.72	2.13
14_2	0.95	2.13	1.17
2_VIEGR2_RN_2	3.58	2.38	1.20
1_INDIACHAB_RN_1	3.89	2.83	1.06
3_INDIACHAB_RN_2	4.51	3.20	1.31
2_INDIACHAB_RN_1	4.79	3.75	1.04
13_2	2.22	4.78	2.56
3_KAAP_RL_1	4.42	4.85	0.43
1_BRAGOU_AN_2	4.28	6.10	1.82
46_2	2.61	6.13	3.51
31_2	3.32	6.60	3.28
3_JAV_RL_1	4.44	6.88	2.44
2_JAV_RL_2	4.32	7.03	2.71

Model Performance Parameters

$\tilde{R}^2$	0.467
Q	0.302
Median of Error in Prediction	1.185

Figure 2APX Prediction model for the aroma attributes studied, as example of the goodness of the aroma variable selected, to define the blueprint of the different aroma notes. A) parameters used to build the model, B) regression curve and validation set fit, C) results of sensory score prediction on the test samples for attributes: a) Bitter, b) Spicy, c) Acid, d) Fruity, e) Flowery, f) Nutty