Chemometric Modeling of Coffee Sensory Notes Through Their Chemical Signatures: Potential and Limits in Defining An Analytical Tool for Quality Control

Chemometric Modeling Of Coffee Sensory Notes Through Their Chemical Signatures: Potential And Limits In Defining An Analytical Tool For Quality Control

Davide Bressanello[†] Erica Liberto[†*], Chiara Cordero[†], Barbara Sgorbini[†], Patrizia Rubiolo[†], Gloria Pellegrino[‡], Manuela R. Ruosi[‡], Carlo Bicchi[†]

[†] Dipartimento di Scienza e Tecnologia del Farmaco, Università degli Studi di Torino, Via Pietro Giuria 9, 10125 Torino, Italy

[‡] Lavazza S.p.A., Strada Settimo 410, 10156 Torino, Italy

*Corresponding author:

Tel: +39 011 6707662,

Fax: +39 011 6707687;

e-mail: erica.liberto@unito.it

ABSTRACT

Aroma is a primary hedonic aspect of a good coffee. Coffee aroma quality is generally defined by cup tasting, which however is time-consuming in terms of panel training and alignment, and too subjective. It is challenging to define a relationship between chemical profile and aroma sensory impact, but it might provide an objective evaluation of industrial products. This study aimed to define the chemical signature of coffee sensory notes, to develop prediction models based on analytical measurements for use at the control level. In particular, the sensory profile was linked with the chemical composition defined by HS-SPME-GC-MS, using a chemometric-driven approach. The strategy was found to be discriminative and informative, identifying aroma compounds characteristic of the selected sensory notes. The predictive ability in defining the sensory scores of each aroma note was used as a validation tool for the chemical signatures characterized. The most reliable models were those obtained for woody, bitter, and acidic properties, whose selected volatiles reliably represented the sensory note fingerprints. Prediction models could be exploited in quality control, but compromises must be determined if they are to become complementary to panel tasting.

Keywords: coffee aroma, sensory note fingerprints, HS-SPME-GC-MS, chemometrics

INTRODUCTION

Coffee aroma and flavor are the most important properties determining consumers' preference and acceptance and, together with price, are the main aspects underlying a coffee's commercial success.[1–3] Coffee companies must evaluate the quality of the beans they aim to buy, to determine the best coffees to use in their blends, and/or which coffees are of sufficiently high quality to be marketed as "single origin". Several attributes are considered in qualifying a good coffee, including lack of defects, bean color and size, and flavor, the latter aspect being 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; they are also rather difficult to apply at-line for immediate feedback. The sensory approach adopted to date 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 flavor attributes quantitatively, through scaled scores. The lexicon used when measuring the sensory aspects of different flavor attributes is therefore of great importance, and must carefully be defined when comparing chemical and sensory data, 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.[6–8] 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 the industrial fields, because it may enable food industries to obtain a further objective evaluation, independent of or complementary to the panel's 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.[9–12] Conversely, several different sampling approaches have been used to extract and concentrate flavor components, and more particularly volatile compounds, directly from ground coffee (powder) and/or from the coffee brew. Sample preparation is the crucial step in any analytical procedure, and must enable the recovery of chemical components representative of each sensory informative note.[1,13–24] Furthermore, in quality control, the entire analytical procedure used to study chemicals related to the sensory experience should be integrated and fully automated. In this context, chemometrics 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 evaluating their probable market impact, and in predicting preferences based on formulation changes.[25,26] Furthermore, chemometrics can provide information about the chemicals that discriminate among sensory attributes, and link those chemicals to sensory perception through correlation models. The conventional strategy in aroma studies implies that a single odorant is considered together with its sensory description. Conversely, correlation through chemometrics makes it possible: to simultaneously measure all compounds eliciting a peculiar sensory perception, and 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.[13,21,27–29] However, reliable models that describe a sensory note representative of the variability of coffee require: a large number of different coffee samples, and fast analytical techniques applicable at-line or on-line, as a complement to the verdict of the sensory panel. Several studies have addressed the sensory-instrumental relationship relating to the sensory properties of coffee, but most looked at correlations among small and uniform pools of samples, without taking into account the wide variability of coffees caused by origin, post-harvest processing, and roasting.[21,30,31]

This study is part of a wider project aiming to correlate the sensory characteristics of coffee aroma with its chemical composition, in order to provide an objective tool that is complementary to sensory evaluation, and that can be applied for routine use.

This part of the study, in particular, focuses on the use of chemometrics as a tool to define the odorants characterizing the chemical signature of different coffee aroma notes, and to validate it for sensory score prediction (Figure 1). Coffee samples with particular sensory characteristics were included as representative of commercial coffees; samples were of different origins, species, and submitted to different post-harvest treatments, as occurs in quality control at the industrial level upon acceptance of incoming beans. Specimens were analyzed 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 combination may also be included within automatic Total Analysis Systems (TAS), with which a large number of samples can be screened for quality control of in-cup coffee sensory quality.[9,32,33] The demand for sensory quality control and evaluation is becoming crucial for coffee producers; the choice of this TAS method aimed to reconcile the need for full characterization with that of screening increasingly large numbers of samples. The choice of analytical strategy was driven by the need to balance these two requirements, but was also responsible for the strategy's potentials and limits.

Materials and Methods

Reagents and Matrices

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

156 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 analyzed in replicates. Table 1 lists twenty samples with their sensorial characterization. Samples 1-8 (five replicates each, *n*=40) were used during the first part of the study to determine the most suitable sample preparation method to study the relationship between chemical and sensory attributes.[34] Samples 9 to 20 (three replicates each, *n*=36) 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; these are also indicated in the text as "stressed samples". The roasting degree of each sample was carefully measured by ground bean light reflectance, with a single-beam Color Test 2 instrument Neuhaus Neotec (Genderkesee, Germany) at a wavelength of 900 nm, on 25-30 g 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 h prior to cupping, and left for at least 8h to stabilize. For clarity, samples are labeled in the text with their origins.

Pure reference standards for identity confirmation were from Sigma-Aldrich (Milan, Italy): table, and *n*-alkanes (n-C9 to n-C25) for Linear Retention Index (LRI) determination.

Internal standards (ISTDs) for analyte response normalization were *n*-C_{13}. 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.

Headspace solid phase microextraction (HS-SPME) sampling

The fiber was 1 cm length, coated with a polydimethylsiloxane/pivinylbenzene (PDMS/DVB), 65 μm, conditioned before use as recommended by the manufacturer. The SPME device was from Supelco (Bellfonte, PA). Coffee aroma compounds were sampled by automated headspace solid phase microextraction (auto-HS-SPME), using a Combi-PAL AOC 5000 (Shimadzu, Milan, Italy),

assembled in-line with a Shimadzu QP2010 GC–MS system, controlled by 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 min at 50 °C at a stirring speed of 350 rpm. The internal standard had previously been loaded onto the fiber[35] by sampling 5µL of a 1000 mg/L solution of *n*-C_{13} in dibuthylphtalate in a 20mL headspace vial for 20 min at 50 °C, stirring speed 350 rpm. 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 analyzed in duplicate.

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) 300 mm x 0.25 mm i.d., 0.25 µm (SGE-Melbourne, Australia). Temperature program, from 40 °C (1 min) to 200 °C at 3 °C/min and to 250 °C at 10 °C/min (5 min).

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

***Identification of volatile components*.** Aroma compounds sampled from the headspace of the coffee powder were identified through their linear retention indices (LRI)[36,37] 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).

Descriptive sensory analysis of coffee aroma

The samples were submitted to sensory evaluation through quantitative descriptive analysis (QDA) by the Lavazza trained panel of experts in coffee evaluation, following the SCAA Q cupping and grading protocol.[4,5,38] This protocol entails three tasting steps, after roasting to a set color (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 3 minutes after its preparation and stirring, and iii) flavor evaluation after 8-10 min. The attributes aftertaste, acidity, body, and balance are evaluated by tasting the brew, spraying it into the mouth to maximize retro-nasal vapors.

In the preliminary part of the study, five panelists determined the notes to be considered: flavor and body (mouthfeel), astringency, and aroma intensity. In the second part, ten panelists assessed cup quality for flowery, fruity, woody, nutty, spicy, acidity, and bitterness. The quality and intensity of each attribute were evaluated simultaneously, upon a scale from 0 to 10. ANOVA analysis was run to verify panel alignment on each attribute. Average scores from experts whose evaluations were similar were used as "main scores" for the attributes under investigation.

Data processing

Chromatographic data were collected using Shimadzu GCMS Solution 2.5SU1 software(Shimadzu). 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 (i.e. analytes target ion areas *versus* the internal standard target ion area). Auto-scaling was applied as data pretreatment: this step ensured that the contribution of each X variable (odorant) to the Y variable (sensory score) was unbiased. One-way ANOVA and t-test on the sensorial results, and PCA, PLS-DA, and PLS, were run by XLSTAT software (Addinsoft, Paris, France).

RESULTS AND DISCUSSION

Qualitative descriptive profiles

PCA is relevant in chemometrics mainly as a standard tool to display the qualitative aroma profiles of samples. In QDA, a panel of trained assessors rates a number of samples for perceived intensities of distinct attributes on scales; the panel follows reference protocols for specific food commodities (e.g. coffee or olive oil)**,** depending on the panel's experience and/or on the complexity of the matrix. By averaging these intensity ratings and replicates, a data matrix may be built up, in which the rows are food samples, and the columns the relative sensory attributes used to describe them[25,26]. Analysis of this data matrix by PCA can give information on both how coffee samples are related, and which sensory notes best describe each sample. PCA was applied to the mean QDA sensory scores for aroma and flavor of eight samples, analyzed in five replicates by five panelists; the bi-plot of scores and loadings are 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 there were correlations between volatiles and taste sensory attributes, since several volatiles and non-volatiles have common reaction pathways during roasting.[21,28,30]

Moreover, panelists perceive odors via ortho- and retro-nasal pathways. This is the result both of compound mutual interactions and of cross-modal effects between odorants and taste, which can amplify or modify perception that, physiologically, occurs in the brain. These interactions do not occur at the molecular level.[39–41]

Bitter notes were also closely related to nutty and astringent notes. In contrast, vectors for fruity, flowery, and acid were 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

221 India sample (INDIA ARAB CHERRY) that shows sensory characteristics more similar to Robusta
222 samples.
223
224 ***From sensory evaluation to the related chemicals: a discriminative and informative guide***
225 A total of 95 compounds were identified (or tentatively identified). 17 compounds were unknown (or
226 not unequivocally identified) by HS-SPME-GC-MS in the coffee powder samples. Table 2 reports
227 the list of identified compounds with their Linear Retention Indices (LRIs). The coffee aroma
228 chemical profile of the first 40 samples obtained by HS-SPME-GC-MS was processed by PCA,
229 together with their sensory scores, to determine: whether groups and/or outliers were present, and the
230 relationship between samples and chemical-sensory variables. As expected, as well as sensory quality
231 analysis, PCA on the aroma chemical profile showed a discrimination ability, driven firstly by species
232 (Figure 3A), and then, within species, by the sensory characteristics peculiar to each origin (Figure 3
233 C, 3E). The Loadings Plot (Figure 3B) clearly showed that the sensory notes were split into two
234 groups: 1) the acid, flowery, fruity notes, which were located in the 1st quadrant of the Cartesian
235 plane, i.e. the location of the Arabica samples; 2) the bitter, nutty, woody, and spicy notes, lying in
236 the 3rd quadrant, i.e. the location of the Robusta samples. Several chemical variables described this
237 sample distribution, and thus these sensory notes (Figure 3B). Several pyrazines, (e.g. 2-*n*-
238 propylpyrazine, 2,6-diethylpyrazine, 2-methyl-3,5-diethylpyrazine, isopropenylpyrazine) and
239 phenolic derivatives, such as guaiacols, characterized the Robusta samples and were more closely
240 related to the roasty, tobacco, nutty, spicy, and woody notes, while furan derivatives, esters and
241 ketones were linked to the sweet, fruity, and floral sensory attributes.[9,22,42]
242 Within the Robusta samples: a) JAVA was the most nutty sample, characterized by compound #48
243 (unk 4); b) UGA samples had high acidity and were chemically described by 2,3-butandione and 2,3-
244 pentandione, acetoxyacetone, hexanal, acetic acid, 1-hydroxy-2-butanone, and 1-*H*-pyrrole-2-
245 carboxaldehyde, compounds elicit musty, sour, pungent, buttery notes that can be related to acid
246 attributes; c) INDO was the most woody, spicy, and bitter sample, and was more full-bodied and

247 astringent, mostly characterized by several pyrazines and phenolic compounds, as mentioned above
248 (Figure 3D).
249 Within Arabica, separation between samples was achieved on PC1 and PC2: a) KAFAs were
250 characterized by high body, fruity note, and aroma intensity; chemical variables related to these
251 sensory attributes were furfuryl alcohol, methyl acetate, 5-methyl furfural, 2-cyclopenten-1-one-3
252 methyl, all of which are characterized by sweet, fruity, malty, and nutty notes; b) conversely, PNG
253 was mostly characterized by acetyl furan, 2-furfuryl-5-methylfurane, 2-furanmethanol propanoate, 2-
254 furfuryl furan (Figure 3F).

255 Although HS-SPME discriminates among analytes depending on their volatility/polarity, if used
256 under standardized conditions it provides reliable information for fingerprinting studies, and is
257 perfectly suitable for comparative analyses.[43] Although PCA applied to comprehensive sensory and
258 chemical data possesses great informative potential, it is difficult to define the chemical fingerprint
259 of a single, specific, note for the purpose of discriminating samples by their sensory characteristics.
260 It is therefore necessary to analyze the relationship between chemical compounds and sensory note
261 descriptions in greater depth, looking at the chemical variables most closely correlated with each
262 sensory note. Chemical variables showing a Pearson correlation coefficient above 0.5 (taken as cut-
263 off) were singled out to compare samples characterized by different sensory notes.

264 Table 2 reports the variables related to each sensory attribute (except for aroma intensity, body, and
265 astringency). It may be seen that specific sensory notes (e.g. acid and flowery, or woody, bitter, nutty,
266 and spicy) are often described statistically by the same variables, i.e. the components statistically
267 correlated with these notes are very often the same. These results confirm that a specific sensory note
268 is described by component amounts and ratios, and rarely by single specific compounds. [17]

269 This may be due to the complexity of odor and taste perceptions, and to their mutual influence on the
270 actual perceived flavor. Interactions among odorants give odor synesthesia, while interactions
271 between odorants and tastants might give chemesthetic perception.[2,3] Several examples of this

possible interaction are reported in Table 2; some note-related compounds were related statistically to the note (X), although their odor description reported in the literature is different when considered alone (X-L). For example, 2,3-Dimethylpyrazine alone has an odor description as nutty, and is also related to other sensory attributes.[39]

The percentage contribution of each compound to the whole chemical profile, correlated statistically to a given sensory attribute, was monitored across all samples investigated. The percentage of each compound was calculated considering their normalized responses (analytes target ion areas *versus* the internal standard target ion area) within the whole targeted profile. Figure 4 shows a "heat map" of the samples, scored on the normalized percentage contributions of components 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 yellow (low percentage) to orange (high percentage). For instance, in Figure 4A, 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, showed sensory characteristics similar to Robusta, thus confirming from the chemical standpoint the sensory scores given by the panel (Table 1). Conversely, guaiacol, 1-*H*-pyrrole-2-carboxaldehyde, (5*H*)-5-methyl-6,7-dihydrocyclopentanpyrazine, 2-furanmethanethiol, and difurfuryldisulfide were directly involved in defining the woody note. [17,19]

Not much may be deduced about the compounds linked to the fruity, acid, and flowery notes, because of the similarity of the chemicals involved. However, the variables found to be correlated to these notes agreed with the sensory scores of those samples (Table 1). For instance, fruity related compounds, such as acetic acid and 3-methylbutanoic acid, were massively present in KAFA samples, which achieved the highest score for the fruity note (Figure 4C**).**

It is difficult to determine which compound specifically contributes to a given note, and how it contributes to it, for two reasons. The first is that the chemical definition of a sample's sensory note (i.e. its aroma signature) is linked to its composition, not only qualitatively but also quantitatively,

297 and, in particular, to the ratios between components. The second is the narrow range of the scores of
298 some notes, e.g. for nutty, from 0 to 3. When the range is narrow, seeking odorants that correlate to
299 the sensory note becomes challenging. To overcome these limitations, the range of sensory scores
300 must be maximized by selecting samples with "stressed" sensory notes, i.e. with high and low values.
301 This enables a more precise definition of the aroma compounds involved with the note, so as to verify
302 the method's ability to correlate them with the sensory fingerprint.

303 A selection of "stressed" sensory samples representative of each note considered, independently of
304 species, origin, and post-harvest treatment, were analyzed and the chemical findings related to the
305 sensory scores (Table 1). "Stressed" samples means a panel selection of a new pool of samples with
306 considerable differences of sensory score within a given note.

307

308 ***Definition of note-related compounds (NRC) on representative "stressed" sensory samples***

309 A supervised chemometric tool (PLS-DA) was applied to study note-related compounds. PLS-DA
310 describes samples by calculating new variables that maximize separation between groups, while
311 minimizing variability within groups. Samples with the lowest score (for each target note) were
312 assigned to class 1, while those with the highest score were assigned to class 2. The impact of each
313 compound on the separation of the pool of samples into the two classes (1 and 2) was evaluated by
314 VIP (Variable Impact On Projections). The cut-off was arbitrarily chosen, for each note, as the point
315 at which the VIP values dropped sharply in the histogram. Some points emerged from this selection:
316 a) a single compound can contribute to the score of more than one sensory note. This was pointed out
317 above (non-supervised data elaboration) and was reported by Ribeiro et al. [28] in a study on prediction
318 models of the quality of Arabica coffee beverages;
319 b) the accuracy of variable selection is not the same for all notes. Variables selected to describe the
320 woody note were probably a consequence of clearer definition of the note (in terms both of maximized

scores and of panel alignment on the lexicon) across the group of samples, which led to more precise identification of the most significant note-related compounds.

Validation of NRC in terms of sensory score prediction capability

The compounds identified with PLS-DA for each sensory attribute were used to study the correlation between chemical and sensory data on the whole data set, through a note-prediction model based on the Partial Least Square Regression (PLS) algorithm. Samples were randomly divided into three groups: a training set (131), a validation set (10), and a test set (15). Figure 5 shows the regression parameters, curve and validation set fit for woody note, and score prediction on the external test samples. The results showed correct sample distribution across the calibration interval, indicating the model is reliable for sensory score prediction. Prediction reliability was evaluated through the Standard Deviation Error in Calibration and in Prediction (SDEC and SDEP) from the predicted *vs.* experimental scores, calculated as follows:

$$SDEC\ (P) = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \widehat{y_{i|i}})^2}{n}}$$

Where:

y_i=experimental value, $\widehat{y_{i|i}}$= predicted value, n= number of observations.

The model showed close correlation between odorants selected and sensory scores. The predictive ability was good, i.e. Q^2= 0.754, with a SDEP=1.175.

Similar results, although not as good, were obtained for the other sensory attributes (Table 3**).** The only exception was the nutty note, for which the model showed poor linearity (R^2=0.467), a wide confidence limit (SDEC= 1.646) and very low predictive ability (SDEP= 1.426). This was chiefly

due to the difficulty over the lexicon used to define nutty and, as a consequence, to determine odorants linked to it.

Although to differing extents, the results show good relationships between selected odorants and sensory scores. Average standard error in score prediction was ± 1, cross-validating the link between the compounds selected and the sensory note description. From this perspective, the chemical composition of woody and flowery notes was also investigated by sensomics.[44,45] The preliminary results of this inter-approach validation showed reasonable consistency between chemometrics and sensomics for some of the target compounds. The compounds already confirmed by the sensomic approach are listed in Figure 6.

The chemical structures of the volatiles identified as discriminant for the woody note are reported in Figure 7A. Figure 7B reports the chemical structures of a selection of the volatiles derived from the chemometric approach highly correlated (correlation coefficient >0.7) with those revealed by sensomics. These chemical compounds (A and B) were used in the PLS model prediction of the woody sensory scores. The chemometric-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 are also limited by its ability to cover the entire range of scores of samples under study, i.e. woody *vs.* flowery. Compared to flowery, the robustness of the woody note gives better prediction, thanks to two factors, namely the wide pool of samples, covering the whole sensory score range; and the good panel alignment in woody note lexicon recognition, providing a better estimation of an external test set.[28–30,46]

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

The chemometric-driven approach was found to be discriminative, informative, and predictive in revealing the chemical signature of the different coffee aroma notes. Discriminative, because it was

able to single out samples with peculiar aroma notes, independently of species, post-harvest treatment, and origins. Informative, because it exploited the complementary and simultaneous use of sensory and chemical data to define odorants able to describe the chemistry of aroma notes, stressing the method's strengths and limitations. Predictive, because the panel-coherent sensory score prediction, based on this chemometric approach, confirmed and reinforced the relevance and significance of the volatiles selected by applying this procedure. Despite these positive results, the final goal is still a distant one, because the models need of a sufficient number of samples to cover the wide variability of samples (including seasonality); the accuracy in note definition and scoring must be improved, chemometric data treatment can be optimized to better fit the dynamic range in sensory evaluation. In any case, compromises must be made in applying a statistical model when sensory attributes are all evaluated together.

The chemometric data-driven approach is promising for predicting sensory scores from chemical data, and appears to provide a complementary tool that can contribute to objective sensory evaluation, despite the great variability of coffee samples (origins, species, treatments, qualities) that are present in the day-to-day situation in quality control at the industrial level.

Acknowledgements

This study was supported by Ricerca Finanziata da Università - Fondo per la Ricerca Locale (Ex 60%) Anno 2014.

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.

Supporting information

Table 1S The complete list of coffee samples analyzed in this study with their characteristics

393 This material is available free of charge via the internet at http://pubs.acs.org.

394

395

References

(1) Sunarharum, W. B.; Williams, D. J.; Smyth, H. E. Complexity of Coffee Flavor: A Compositional and Sensory Perspective. *Food Res. Int.* **2014**, *62*, 315–325.

(2) Chiralertpong, A.; Acree, T. E.; Barnard, J.; Siebert, K. J. Taste–Odor Integration in Espresso Coffee. *Chemosens. Percept.* **2008**, *1*, 147–152.

(3) Prescott, J. Multisensory Processes in Flavour Perception and Their Influence on Food Choice. *Curr. Opin. Food Sci.* **2015**, *3*, 47–52.

(4) Britta Folmer. *The craft and science of coffee*; Eds.; Elsevier: London UK, **2017**.

(5) SCAA. Protocols cupping specialty coffee. *Spec. Coffee Assoc. Am.* **2015**, 1–10. http://www.scaa.org/PDF/resources/cupping-protocols.pdf [Accessed: December 2016]

(6) WCR (World Coffee Research). *https://worldcoffeeresearch.org/* [Accessed: November 2017]

(7) Kubota, L. Reinventing the (Flavor) Wheel: industry collaborates to identify coffee flavor. *http://www.scanews.coffee/2016/01/19/reinventing-the-flavor-wheel-industry-collaborates-to-identify-coffee-flavor-attributes/* [Accessed: December 2017]

(8) Chambers, E.; Sanchez, K.; Phan, U. X. T.; Miller, R.; Civille, G. V.; Di Donfrancesco, B. Development of a “Living” Lexicon for Descriptive Sensory Analysis of Brewed Coffee. *J. Sens. Stud.* **2016**, *31*, 465–480.

(9) Liberto, E.; Ruosi, M. R.; Cordero, C.; Rubiolo, P.; Bicchi, C.; Sgorbini, B. Non-Separative Headspace Solid Phase Microextraction-Mass Spectrometry Profile as a Marker to Monitor Coffee Roasting Degree. *J. Agric. Food Chem.* **2013**, *61*, 1652–1660.

(10) Picó, Y. *Chemical Analysis of Food: Techniques and Applications*. Eds.; Elsevier: Waltham MA, USA, **2012**.

(11) Yeretzian, C. Coffee. In *Handbook of Odor*; Buettner, A., Ed.; Springer, **2017**; pp 107–127.

(12) 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.* **2015**, *66*, 505-513.

(13) Blank, I.; Alina, S.; Grosch, W. Potent Odorants of the Roasted Powder and Brew of Arabica Coffee. *Zeitschrift fur Leb. Und-forsch.* **1992**, *195*, 239–245.

(14) Bhumiratana, N.; Adhikari, K.; Chambers, E. Evolution of sensory aroma attributes from coffee beans to brewed coffee. *LWT - Food Sci. Technol.* **2011**, *44*, 2185–2192.

(15) Czerny, M.; Grosch, W. Potent odorants of raw arabica coffee. their changes during roasting. *J. Agric. Food Chem.* **2000**, *48*, 868–872.

(16) Czerny, M.; Schieberle, P. Changes in roasted coffee aroma during storage - influence of the packaging. *ASIC-19eme Colloq. Sci. Int. sur le Cafè* **2001**, 1–6.

(17) Folmer, B. How can science help to create new value in coffee? *Food Res. Int.* **2014**, *63*, 477–482.

(18) Flament, I. *Coffee flavor chemistry*; Eds.; Wiley: Chichester UK, **2002**.

(19) Frank, O.; Zehentbauer, G.; Hofmann, T. Bioresponse-guided decomposition of roast coffee beverage and identification of key bitter taste compounds. *Eur. Food Res. Technol.* **2006**, *222*, 492–508.

(20) Jiang, D.; Peterson, D. G. Role of hydroxycinnamic acids in food flavor: a brief overview. *Phytochem. Rev.* **2010**, *9*, 187–193.

(21) Lindinger, C.; Labbe, D.; Pollien, P.; Rytz, A.; Juillerat, M.A.; Yeretzian, C.; Blank, I. When machine tastes coffee: instrumental approach to predict the sensory profile of espresso coffee. Anal. Chem. 2008, 80, 1574–1581.

(22) Ruosi, M. R.; Cordero, C.; Cagliero, C.; Rubiolo, P.; Bicchi, C.; Sgorbini, B.; Liberto, E. A further tool to monitor the coffee roasting process: aroma composition and chemical indices. J. Agric. Food Chem. 2012, 60, 11283–11291.

(23) Scheidig, C.; Czerny, M.; Schieberle, P. Changes in key odorants of raw coffee beans during storage under defined conditions. *J. Agric. Food Chem.* **2007**, *55*, 5768–5775.

(24) Caprioli, G.; Cortese, M.; Cristalli, G.; Maggi, F.; Odello, L.; Ricciutelli, M.; Sagratini, G.;

448 Sirocchi, V.; Tomassoni, G.; Vittori, S. Optimization of espresso machine parameters through
449 the analysis of coffee odorants by HS-SPME-GC/MS. *Food Chem.* **2012**, *135*, 1127–1133.
450 (25) Ochi, H.; Sumi, M.; Nakata, I.; Saito, H.; Uozumi, M.; Iwatsuki, K. Sensometric calibration of
451 sensory characteristics of commercially available milk products with instrumental data. *J.*
452 *Dairy Sci.* **2010**, *93*, 1794–1806.
453 (26) Tomic, O. Using Sensometrics to Extract Information from Your Sensory Data. *Food Sci.*
454 *Technol.* **2012**, *26*, 42–44.
455 (27) Semmelroch, P.; Grosch, W. Studies on Character Impact Odorants of Coffee Brews. *J Agric*
456 *Food Chem* **1996**, *44*, 537–543.
457 (28) Ribeiro, J. S.; Augusto, F.; Salva, T. J. G.; Ferreira, M. M. C. Prediction Models for Arabica
458 Coffee Beverage Quality Based on Aroma Analyses and Chemometrics. *Talanta* **2012**, *101*,
459 253–260.
460 (29) Pongsuwan, W.; Fukusaki, E.; Bamba, T.; Yonetani, T.; Yamahara, T.; Kobayashi, A.
461 Prediction of Japanese Green Tea Ranking by Gas Chromatography/Mass Spectrometry-Based
462 Hydrophilic Metabolite Fingerprinting. *J. Agric. Food Chem.* **2007**, *55*, 231–236.
463 (30) Ribeiro, J. S.; Augusto, F.; Salva, T. J. G.; Thomaziello, R. A.; Ferreira, M. M. C. Prediction
464 of Sensory Properties of Brazilian Arabica Roasted Coffees by Headspace Solid Phase
465 Microextraction-Gas Chromatography and Partial Least Squares. *Anal. Chim. Acta* **2009**, *634*
466 , 172–179.
467 (31) Ribeiro, J. S.; Ferreira, M. M. C.; Salva, T. J. G. Chemometric Models for the Quantitative
468 Descriptive Sensory Analysis of Arabica Coffee Beverages Using near Infrared Spectroscopy.
469 *Talanta* **2011**, *83*, 1352–1358.
470 (32) Cagliero, C.; Bicchi, C.; Cordero, C.; Rubiolo, P.; Sgorbini, B.; Liberto, E. Fast Headspace-
471 Enantioselective GC–mass Spectrometric-Multivariate Statistical Method for Routine
472 Authentication of Flavoured Fruit Foods. *Food Chem.* **2012**, *132*, 1071–1079.
473 (33) Manz, A.; Graber, N.; Widmer, H. M. Miniaturized Total Chemical Analysis Systems: A

474 Novel Concept for Chemical Sensing. *Sensors Actuators B Chem.* **1990**, *1*, 244–248.
475 (34) Bressanello, D.; Liberto, E.; Cordero, C.; Rubiolo, P.; Pellegrino, G.; Ruosi, M. R.; Bicchi, C.
476 Coffee Aroma: Chemometric Comparison of the Chemical Information Provided by Three
477 Different Samplings Combined with GC–MS to Describe the Sensory Properties in Cup. *Food*
478 *Chem.* **2017**, *214*, 218–226.
479 (35) Wang, Y.; O'Reilly, J.; Chen, Y.; Pawliszyn, J. Equilibrium In-Fibre Standardisation
480 Technique for Solid-Phase Microextraction. *J. Chromatogr. A* **2005**, *1072*, 13–17.
481 (36) National Institute of Standard and Technology. Nist Chemistry WebBook
482 https://webbook.nist.gov/chemistry/name-ser/. [Accessed: September 2016]
483 (37) The Good Scents Company (tgsc). The Good Scients Company Information System
484 http://www.thegoodscentscompany.com/. [Accessed: October 2016]
485 (38) Lawless, H. T.; Heymann, H. *Sensory Evaluation of Food*; Food Science Text Series; Springer
486 New York: New York, NY, **2010**.
487 (39) Chambers IV, E.; Koppel, K. Associations of Volatile Compounds with Sensory Aroma and
488 Flavor: The Complex Nature of Flavor. *Molecules* **2013**, *18*, 4887–4905.
489 (40) Donishi; Kimura, A.; Imbe, H.; Yokoi, I.; Kaneoke, Y. Sub-Threshold Cross-Modal Sensory
490 Interaction in the Thalamus: Lemniscal Auditory Response in the Medial Geniculate Nucleus
491 Is Modulated by Somatosensory Stimulation. *Neuroscience* **2011**, *174*, 200–215.
492 (41) Dalton, P.; Doolittle, N.; Nagata, H.; Breslin, P. a. The Merging of the Senses: Integration of
493 Subthreshold Taste and Smell. *Nat. Neurosci.* **2000**, *3*, 431–432.
494 (42) Grosch, W. Flavour of Coffee. A Review. *Nahrung / Food* **1998**, *42*, 344–350.
495 (43) Bicchi, C.; Cordero, C.; Liberto, E.; Sgorbini, B.; Rubiolo, P. Headspace Sampling in Flavor
496 and Fragrance Field. In *Comprehensive Sampling and Sample Preparation*; Elsevier, **2012**;
497 Vol. 4, pp 1–25.
498 (44) Schieberle, P. New Developments in Methods for Analysis of Volatile Flavor Compounds and
499 Their Precursors. In *Characterization of Food*; Elsevier, **1995**; pp 403–431.

(45) 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.*; Henryk, J., Ed.; **2011**; pp 413–438.

(46) 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.* **2007**, *55*, 9908–9912.

507 **Figure captions**

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

509 Figure 2 Biplot of scores and factor loadings obtained by PCA for aroma descriptors of QDA for
510 coffees 1-8 (*n*=40). BRA: Brazil; COL: Colombia; JAV: Java; UGA: Uganda; PNG: Papua New
511 Guinea; INDO: Indonesia.

512 Figure 3 Scores and loading plots of the coffee samples obtained by chemical and sensory analysis,
513 A and B all samples, C and D Robusta, E and F Arabica coffees. BRA: □; COL: ◊; JAV: Δ; UGA: X;
514 PNG: * ; INDIA: ▭ ; INDO: ◯ ; KAFA: +. For sample acronym see Table 1 and for loading
515 identification see Table 2.

516 Figure 4 Heat map of the percentage contribution of compounds correlated with: A) woody, B) nutty
517 and C) fresh notes (acid, flowery, fruity)

518 Figure 5 A) parameters used to build the model, B) regression curve and validation set fit for the
519 woody note, C) results of sensory score prediction on an external test samples

520 Figure 6 Compounds characterizing woody and flowery aroma notes confirmed by sensomics.

521 Figure 7 Chemical structures of the compounds highlighted by molecular sensory science as
522 important in the characterization of woody and used in the woody note chemometric prediction
523 model. (A) compounds shared between the two approaches (B) compounds derived from the
524 chemometric-driven approach as highly correlated (correlation coefficient >0.7) with those pointed
525 out by sensomics (Figure 6). Both (A) and (B) compounds were used in the PLS model prediction of
526 the woody sensory scores.

527

528 Table 1 Coffee samples and sensory characteristics. Species (A: Arabica, R: Robusta), Treatment (N: Natural, W: Washed).

#	Sample acronym	Type	Species	Treatment	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	-	-

529

Table 2 List of identified and *tentatively identified compounds.

#	Compound Name	Odor Description[11,36,37]	Calc. LRI	Lit. LRI	Acid	Fruity	Flowery	Bitter	Nutty	Woody	Spicy
1	Acetaldehyde	*Fruity*[11]	706	723							
2	Acetone	*Ethereal, Apple, Pear*[37]	824	835							
3	Methyl acetate	*Ethereal, Sweet, Fruit*[37]	828	839							
4	2-Methylfuran	*Chocolate, burnt, ethereal acetate*[11]	885	864							
5	2-Butanone	*Ethereal*[37]	906	905							
6	2-Methylbutanal	*Chocolate*[11]	914	931							
7	3-Methylbutanal	*Malty*[11]	918	936							
8	2,5-Dimethylfuran	*Ethereal*[37]	950	939							
9	2,3-Butanedione	*Buttery*[11]	978	963							
10	2,3-Pentanedione	*Buttery*[11]	1058	1060	X		X				
11	2-Vinylfuran	*Nutty, coffee*[37]	1071	1085							
12	Hexanal	*Tallowy, leaf-like*[11]	1080	1098							
13	2,3-Hexanedione	*Buttery*[37]	1130	1110							
14	1-Methyl-1*H*-pyrrole	*Woody*[37]	1137	1140							
15	2-Vinyl-5-methylfuran	-	1151	1152							
16	Pyridine	*Fishy*[11]	1177	1177							
17	Pyrazine	*Sweet, Floral*[37]	1209	1206				X			
18	Methylpyrazine	*Nutty*[37]	1262	1268				X			
19	3-Hydroxy-2-butanone	*Buttery*[37]	1281	1285							
20	1-Hydroxy-2-propanone	*Sweet-Caramel-like*[37]	1297	1318	X		X				
21	2,5-Dimethylpyrazine	*Nutty-Roasted, Cocoa*[37]	1317	1321				X		X	X
22	2,6-Dimethylpyrazine	*Earthy-Chocolaty*[37]	1324	1327				X		X	X
23	Ethylpyrazine	*Nutty-roasted*[37]	1329	1343				X	X-L	X	X
24	2,3-Dimethylpyrazine	*Nutty, coffee, peanut butter, walnut, caramel eather*[37]	1341	1354				X	X-L	X	X
25	1-Hydroxy-2-butanone	*Sweet-caramel-like*[36]	1370	1381	X		X				

#	Compound Name	Odor Description[11,36,37]	Calc. LRI	Lit. LRI	Acid	Fruity	Flowery	Bitter	Nutty	Woody	Spicy
26	3-Ethylpyridine	*Tobacco, caramel-like*[37]	1374	1384				X		X	X
27	2-Ethyl-5-methylpyrazine	*Coffee-like*[37]	1386	1399				X	X-L	X	X
28	2-Ethyl-3-methylpyrazine+ Trimethylpyrazine	*Raw Potato*[37]	1399	1400				X-L	X	X	X
29	Unk 1(m/z: 54 [100%]; 43[78%]; 42[16.75%])	-	1405	-	X		X				
30	2-*n*-Propylpyrazine *	*Green vegetable*[11]	1413	1425				X	X	X	X-L
31	Unk 2(m/z: 112 [100%]; 68[73.76%]; 40[24.93%])	-	1418	-	X		X				
32	2-Furanmethanethiol	*Roasted, Burnt, coffee-like*[37]	1432	1440				X	X	X-L	X
33	2-Ethyl-3,6-dimethylpyrazine	*Earthy, Baked*[11]	1441	1449				X-L	X	X	X
34	Acetic acid	*Sour, Pungent*[11]	1446	1454	X-L	X					
35	2,6-Diethylpyrazine	*Hazelnut-like*[37]	1457	1458				X	X-L	X	X
36	Furfural	*Sweet, Woody, Bready*[11]	1462	1467	X		X				
37	Acetoxyacetone	*Fruity, Berry*[37]	1468	1467							
38	Furfurylmethylsulfide	*Coffee-like odor*[37]	1483	1496				X-L		X	X
39	3,5-Diethyl-2-methylpyrazine	*Coffee-like*[37]	1489	1491				X	X	X	X
40	2,5-Dimethyl-3(2*H*)-furanone	*Caramel-like, fruity*[37]	1496	1490	X		X				
41	Furfuryl formate	*Ethereal*[11]	1497	1497							
42	Acetylfuran	*Sweet-caramel-like*[37]	1500	1498	X		X				
43	3-Methyl-2-Cyclopenten-1-one + 3,5-Diethyl-2-Methylpyrazine	*Burnt, Rubbery, earthy + coffee-like*[37]	1509	1509				X	X	X	X
44	1-H-Pyrrole	*Sweet:ethereal*[11]	1513	1525				X	X	X	X-L
45	Benzaldehyde + 2-Methyl-3(2*H*)-thiophenone *+ Furan-2-yl-propan-2-ol	*Sweet*[37]	1519	1520							
46	2-Oxopropylpropanoate	-	1531	1531							
47	Furfuryl acetate	*Garlic, pungent vegetable, onion*[11]	1538	1539							
48	Unk 4(m/z: 110[100%]; 109[86.72%]; 53[50.99%])	-	1564	-							
49	5-Methyl Furfural	*Caramel-like*[37]	1570	1562	X		X				
50	2,3-Butandiole + 1-(5-Methyl-2-furyl)2-propanone	-	1575	1582							
51	1-Methylethenylpyrazine	*Roasted, Nutty*[36]	1590	-				X	X-L	X	X

#	Compound Name	Odor Description[11,36,37]	Calc. LRI	Lit. LRI	Acid	Fruity	Flowery	Bitter	Nutty	Woody	Spicy
52	Furfuryl propanoate	*Fruity*[37]	1598	1603							
53	2-Furfurylfuran	*Caramel-like, earthy, mushroom*[37]	1608	-							X
54	(5*H*)-5-Methyl-6,7-dihydrocyclopentapyrazine	*Earthy*[36]	1611	1611				X-L	X	X	X
55	1-Methylpyrrole-2-carboxaldehyde	*Cracked/pop-corn*[37]	1614	1635							
56	4-Hydroxybutanoate	-	1621	-							
57	Unk 6 (m/z: 137[100%]; 94[61.57%]; 122[37.71%])	-	1630	-							
58	2-Isopropenylpyrazine	*Caramel-like, Nutty*[36]	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	*Mild, slighly caramel-like*[37]	1661	1664							
62	3-Methylbutanoic acid	*Acid, Herbaceous, Sour*[37]	1667	1670	X-L	X-L	X-L				
63	3-Methyl-1,2-ciclohexanedione	*Saffron, burnt, chemical*[37]	1678	-							
64	2-Furfuryl-5-methylfuran	*Alliaceous, earthy, mushroom*[37]	1686	1636							
65	2-Acetyl-3-methylpyrazine*	*Nutty*[36]	1694	1719				X	X-L	X	X
66	Furfurylpentanoate	*Fruity*[36]	1702	1702							
67	2-Methyl-6-(1-propenyl)-pyrazine*	-	1708	1719				X	X	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	-							
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

#	Compound Name	Odor Description[11,36,37]	Calc. LRI	Lit. LRI	Acid	Fruity	Flowery	Bitter	Nutty	Woody	Spicy
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*[36]	1786	-							
79	Unk 20 (m/z:139; 43[12.25%]; 154[50.48%])	-	1791	-				X			
80	2-Hydroxy-3-methyl-2-cyclopenten-1-one	*Caramel-like-Spicy*[37]	1820	1839							
81	Furfurylpyrrole	*Geen, hay-like*[37]	1853	1866				X		X	X
82	Guaiacol	*Smoky/sweet*[11]	1860	-				X		X	X-L
83	2-Acetyl-5-methylfuran	*Caramel-like*[37]	1885	-	X		X				
84	3-Ethyl-2-hydroxy-2-Cyclopenten-1-one	*Caramel-like*[37]	1885	-							
85	*trans*-Furfurylideneacetone*	-	1897	-							
86	Phenylethanol	*Floral, woody, honey-like*[37]	1902	1912				X		X-L	X
87	Maltol	*Sweet, Caramel-like*[37]	1952	1960							
88	2-Acetylpyrrole	*Bitter, Roasted*[37]	1962	1971				X			X
89	Difurfurylether	*Coffee-like, mushroom*[37]	1980	1977				X			X
90	Phenol deriv*	-	1997	-				X		X	X
91	1H-Pyrrole-2-carboxaldehyde	*Corny-Pungent*[37]	2012	2035							
92	4-Ethylguaiacol	*Spicy*[11]	2021	2037				X		X	X-L
93	Nonanoic acid	*Nut-like, Fatty*[37]	2150	2159							
94	4-Vinylguaiacol	*Clove-like*[11]	2185	2193				X		X-L	X-L
95	Difurfuryldisulfide	*Mushrooms, caramel-like*[37]	2536	2536							

Compounds with an "X" are related to each sensory note; "L" indicates that the compound alone directly elicits the peculiar note perception. Odor description is taken from literature references 11, 36, 37.

Table 3

Summary of the performances of sensory notes prediction models

AROMA NOTE	$R^2_{LMO(10\%)}$	$Q^2_{(10\%)}$	SDEC	$Q^2_{(15\%)EXT}$	$SDEP_{EXT}$
Woody	0.816	0.659	0.977	0.754	1.175
Bitter	0.806	0.689	0.755	0.830	0.670
Spicy	0.684	0.298	1.337	0.653	1.155
Acid	0.831	0.497	1.019	0.734	0.890
Fruity	0.661	0.193	0.837	0.568	1.115
Flowery	0.793	0.342	1.047	0.475	1.244
Nutty	0.467	0.302	1.646	0.634	1.426

Figure 1

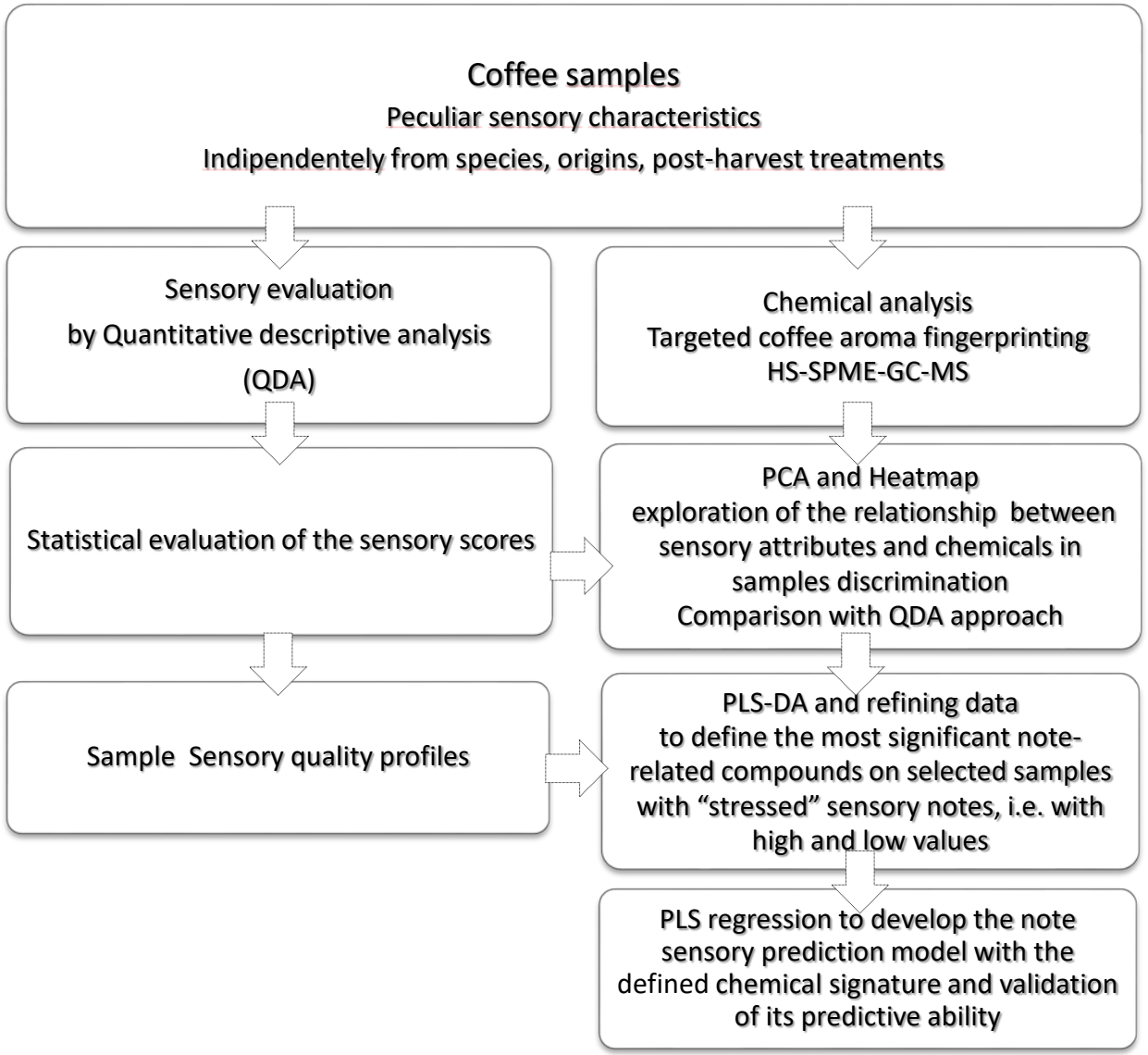

Coffee samples
Peculiar sensory characteristics
Indipendentely from species, origins, post-harvest treatments
Sensory evaluation
by Quantitative descriptive analysis
(QDA)
Chemical analysis
Targeted coffee aroma fingerprinting
HS-SPME-GC-MS
Statistical evaluation of the sensory scores
PCA and Heatmap
exploration of the relationship between sensory attributes and chemicals in samples discrimination
Comparison with QDA approach
Sample Sensory quality profiles
PLS-DA and refining data
to define the most significant note-related compounds on selected samples with "stressed" sensory notes, i.e. with high and low values
PLS regression to develop the note sensory prediction model with the defined chemical signature and validation of its predictive ability

Figure 2

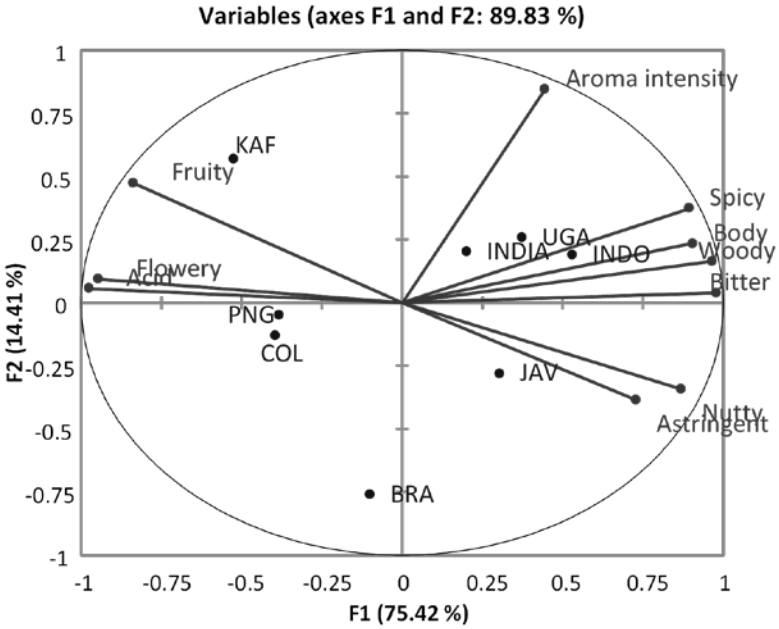

Variables (axes F1 and F2: 89.83 %)
Aroma intensity
KAF
Fruity
Spicy
Body
UGA
INDIA
INDO
Woody
Flowery
Acid
Bitter
PNG
COL
JAV
Nutty
Astringent
BRA
F2 (14.41 %)
F1 (75.42 %)
1
0.75
0.5
0.25
0
-0.25
-0.5
-0.75
-1
-1
-0.75
-0.5
-0.25
0
0.25
0.5
0.75
1

Figure 3

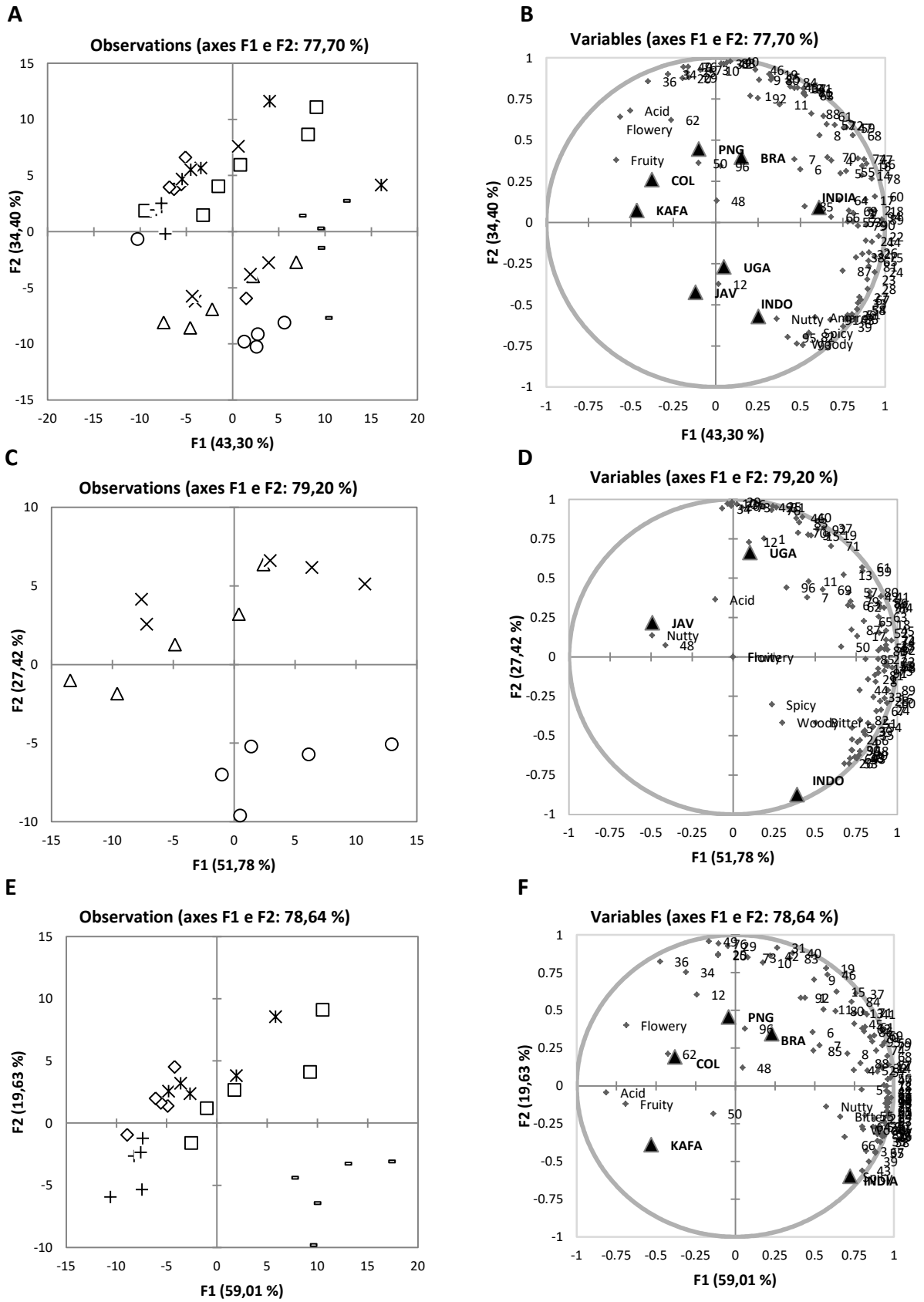

A
Observations (axes F1 e F2: 77,70 %)
F2 (34,40 %)
F1 (43,30 %)
B
Variables (axes F1 e F2: 77,70 %)
Acid
Flowery
Fruity
PNG
BRA
COL
KAFA
INDIA
UGA
JAV
INDO
Nutty
Spicy
Woody
F2 (34,40 %)
F1 (43,30 %)
C
Observations (axes F1 e F2: 79,20 %)
F2 (27,42 %)
F1 (51,78 %)
D
Variables (axes F1 e F2: 79,20 %)
UGA
Acid
JAV
Nutty
Flowery
Spicy
Woody
Bitter
INDO
F2 (27,42 %)
F1 (51,78 %)
E
Observation (axes F1 e F2: 78,64 %)
F2 (19,63 %)
F1 (59,01 %)
F
Variables (axes F1 e F2: 78,64 %)
Flowery
PNG
BRA
COL
Acid
Fruity
Nutty
Bitter
KAFA
INDIA
F2 (19,63 %)
F1 (59,01 %)

Figure 4

A WOODY	BRA	COL	JAV	UGA	PNG	INDIA	INDO	KAFA
2,5-Dimethylpyrazine	2.867	2.665	3.638	3.402	2.539	3.521	3.980	3.453
2,6-Dimethylpyrazine	3.460	3.088	4.515	4.018	3.130	4.298	4.681	3.529
Ethylpyrazine	1.437	1.208	2.220	1.957	1.216	1.779	2.192	1.353
2,3-Dimethylpyrazine	0.637	0.514	0.903	0.823	0.536	0.894	1.078	0.548
3-Ethylpyridine	0.121	0.081	0.119	0.132	0.101	0.198	0.229	0.100
2-Ethyl-5-methylpyrazine	1.260	1.121	2.230	1.888	1.095	1.781	2.256	1.370
2-Ethyl-3-methylpyrazine + trimethylpyrazyne	1.589	1.346	2.640	2.270	1.359	2.458	2.979	1.568
2-*n* -Propylpyrazine	0.048	0.045	0.090	0.070	0.044	0.071	0.109	0.039
2-Furanmethanethiol	0.294	0.215	0.396	0.380	0.237	0.313	0.429	0.181
2-Ethyl-3,6-Dimethylpyrazine	0.907	0.753	1.840	1.500	0.659	1.649	1.991	0.659
2,6-Diethylpyrazine	0.348	0.286	0.726	0.597	0.296	0.619	0.898	0.356
Furfurylmethylsulfide	0.172	0.192	0.204	0.239	0.220	0.301	0.425	0.196
3,5-Diethyl-2-Methylpyrazine	0.152	0.124	0.391	0.312	0.134	0.324	0.454	0.174
3 Methyl-2-Cyclopenten-1-one + 3,5-Diethyl-2-Methylpyrazine	0.036	0.027	0.090	0.081	0.028	0.100	0.137	0.042
1*H* -Pyrrole	0.544	0.345	0.643	0.572	0.397	0.646	0.693	0.332
1-Methylethenylpyrazine	0.148	0.122	0.278	0.224	0.124	0.209	0.318	0.134
(5*H*)-5-Methyl-6,7-Dihydrocyclopentapyrazine	0.204	0.173	0.374	0.326	0.183	0.321	0.456	0.188
1-Methylpyrrole-2-Carboxaldehyde	0.878	0.738	0.820	0.895	0.860	0.944	0.976	1.089
2-Isopropenylpyrazine	0.226	0.170	0.423	0.367	0.186	0.407	0.573	0.195
2-Acetyl-3-Methylpyrazine	0.373	0.314	0.560	0.461	0.315	0.454	0.504	0.343
2-Methyl-6-(1-propenyl)-pyrazine	0.184	0.149	0.355	0.289	0.149	0.337	0.405	0.179
Methyl nicotinate +other	0.097	0.084	0.130	0.119	0.076	0.123	0.138	0.069
Furfurylpyrrole	1.047	1.053	1.666	1.368	1.120	1.315	1.707	1.032
Guaiacol	0.382	0.416	1.010	1.556	0.479	0.739	2.023	0.442
Benzeneethanol	0.159	0.104	0.143	0.319	0.132	0.192	0.270	0.095
Phenol deriv	0.275	0.291	0.533	0.657	0.300	0.402	0.712	0.285
4-Ethyl-guaiacol	0.319	0.302	1.188	1.251	0.302	0.610	2.220	0.173
4-Vinyl-guaiacol	1.828	1.944	5.542	5.350	1.609	1.879	5.397	1.015

B NUTTY	BRA	COL	JAV	UGA	PNG	INDIA	INDO	KAFA
Ethylpyrazine	1.437	1.208	2.220	1.957	1.216	1.779	2.192	1.353
2,3-Dimethylpyrazine	0.637	0.514	0.903	0.823	0.536	0.894	1.078	0.548
2-Ethyl-5-methylpyrazine	1.260	1.121	2.230	1.888	1.095	1.781	2.256	1.370
2-Ethyl-3-methylpyrazine + Trimethylpyrazyne	1.589	1.346	2.640	2.270	1.359	2.458	2.979	1.568
2-*n* -Propylpyrazine	0.048	0.045	0.090	0.070	0.044	0.071	0.109	0.039
2-Furanmethanethiol	0.294	0.215	0.396	0.380	0.237	0.313	0.429	0.181
2-Ethyl-3,6-dimethylpyrazine	0.907	0.753	1.840	1.500	0.659	1.649	1.991	0.659
2,6-Diethylpyrazine	0.348	0.286	0.726	0.597	0.296	0.619	0.898	0.356
3,5-Diethyl-2-methylpyrazine	0.152	0.124	0.391	0.312	0.134	0.324	0.454	0.174
3 Methyl-2-cyclopenten-1-one + 3,5-Diethyl-2-methylpyrazine	0.036	0.027	0.090	0.081	0.028	0.100	0.137	0.042
1*H* -Pyrrole	0.544	0.345	0.643	0.572	0.397	0.646	0.693	0.332
1-Methylethenylpyrazine	0.148	0.122	0.278	0.224	0.124	0.209	0.318	0.134
(5*H*)-5-methyl-6,7-dihydrocyclopentapyrazine	0.204	0.173	0.374	0.326	0.183	0.321	0.456	0.188
2-Isopropenylpyrazine	0.226	0.170	0.423	0.367	0.186	0.407	0.573	0.195
2-Acetyl-3-methylpyrazine	0.373	0.314	0.560	0.461	0.315	0.454	0.504	0.343
2-Methyl-6-(1-propenyl)-pyrazine	0.184	0.149	0.355	0.289	0.149	0.337	0.405	0.179
Methyl nicotinate +other	0.097	0.084	0.130	0.119	0.076	0.123	0.138	0.069
Guaiacol	0.382	0.416	1.010	1.556	0.479	0.739	2.023	0.442
Benzenethanol	0.159	0.104	0.143	0.319	0.132	0.192	0.270	0.095
Phenol deriv	0.275	0.291	0.533	0.657	0.300	0.402	0.712	0.285
4-Ethyl-guaiacol	0.319	0.302	1.188	1.251	0.302	0.610	2.220	0.173
4-Vinyl-guaiacol	1.828	1.944	5.542	5.350	1.609	1.879	5.397	1.015

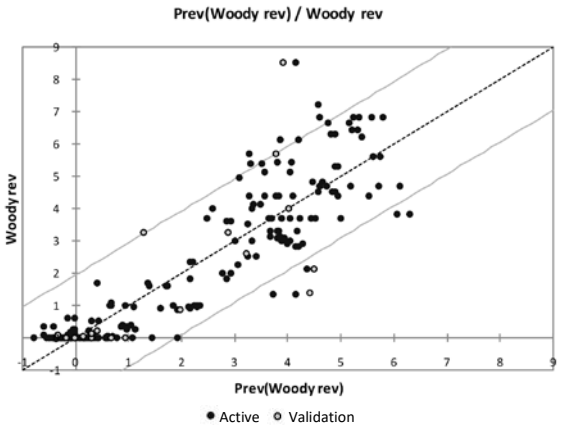
C ACID/ FRUITY / FLOWERY	BRA	COL	JAV	UGA	PNG	INDIA	INDO	KAFA
2,3-Pentanedione	0.299	0.337	0.189	0.184	0.346	0.190	0.094	0.314
1-Hydroxy-2-propanone	0.855	1.123	0.584	0.541	0.924	0.484	0.303	0.768
1-Hydroxy-2-butanone	0.077	0.098	0.044	0.048	0.080	0.042	0.027	0.071
Unk 1	0.379	0.447	0.246	0.241	0.375	0.177	0.110	0.329
Unk 2	0.132	0.143	0.090	0.100	0.140	0.089	0.059	0.138
Acetic acid	4.783	7.639	3.544	3.163	7.246	3.826	2.041	6.848
Furfural	3.243	5.369	2.003	2.334	4.131	1.600	1.020	4.023
2,5-Dimethyl-3(2H)-Furanone	0.645	0.730	0.484	0.521	0.739	0.501	0.357	0.739
Acetylfuran	1.393	1.892	0.926	1.104	1.814	1.202	0.990	1.784
5-Methyl Furfural	6.024	7.797	4.483	5.187	7.154	3.995	2.880	7.350
3-Methylbutanoic acid	2.346	3.072	2.140	2.713	3.533	2.032	2.412	5.151
Unk 14	0.263	0.327	0.192	0.173	0.299	0.189	0.123	0.273
Unk 17	0.338	0.471	0.258	0.293	0.430	0.241	0.182	0.419
Unk 21	0.826	1.032	0.563	0.671	1.068	0.696	0.412	0.937

Figure 5

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

Figure 6

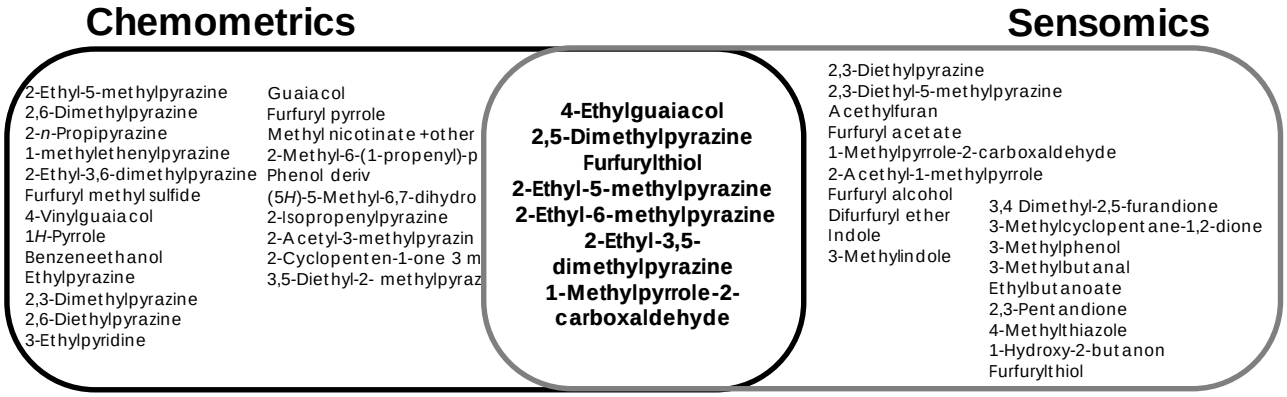
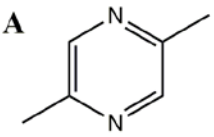
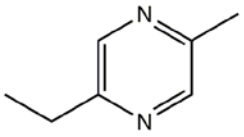

Chemometrics
Sensomics
2-Ethyl-5-methylpyrazine
2,6-Dimethylpyrazine
2-n-Propipyrazine
1-methylethenylpyrazine
2-Ethyl-3,6-dimethylpyrazine
Furfuryl methyl sulfide
4-Vinylguaiacol
1H-Pyrrole
Benzeneethanol
Ethylpyrazine
2,3-Dimethylpyrazine
2,6-Diethylpyrazine
3-Ethylpyridine
Guaiacol
Furfuryl pyrrole
Methyl nicotinate +other
2-Methyl-6-(1-propenyl)-p
Phenol deriv
(5H)-5-Methyl-6,7-dihydro
2-Isopropenylpyrazine
2-Acetyl-3-methylpyrazin
2-Cyclopenten-1-one 3 m
3,5-Diethyl-2- methylpyraz
4-Ethylguaiacol
2,5-Dimethylpyrazine
Furfurylthiol
2-Ethyl-5-methylpyrazine
2-Ethyl-6-methylpyrazine
2-Ethyl-3,5-dimethylpyrazine
1-Methylpyrrole-2-carboxaldehyde
2,3-Diethylpyrazine
2,3-Diethyl-5-methylpyrazine
Acethylfuran
Furfuryl acetate
1-Methylpyrrole-2-carboxaldehyde
2-Acethyl-1-methylpyrrole
Furfuryl alcohol
Difurfuryl ether
Indole
3-Methylindole
3,4 Dimethyl-2,5-furandione
3-Methylcyclopentane-1,2-dione
3-Methylphenol
3-Methylbutanal
Ethylbutanoate
2,3-Pentandione
4-Methylthiazole
1-Hydroxy-2-butanon
Furfurylthiol

Figure 7

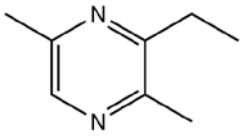
A



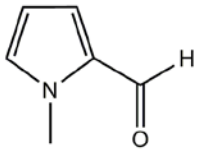
2,5-Dimethylpyrazine, **21**



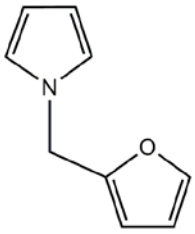
2-Ethyl-5-methylpyrazine, **27**



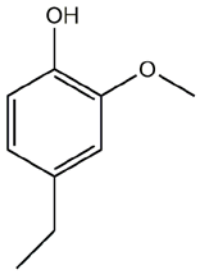
2-Ethyl-3,6-dimethylpyrazine, **33**



1-Methylpyrrole-2-carboxaldehyde, **55**

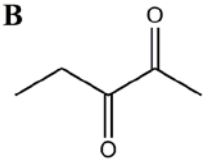


Furfurylpyrrole, **81**

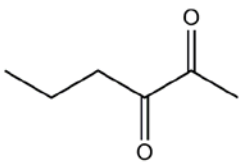


4-Ethylguaiacol, **92**

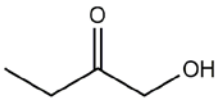
B



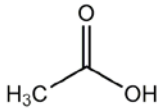
2,3-Pentanedione, **10**



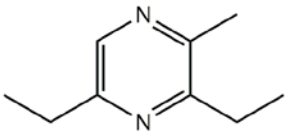
2,3-Hexandione, **13**



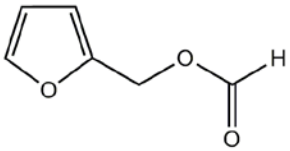
1-Hydroxy-2-butanone, **25**



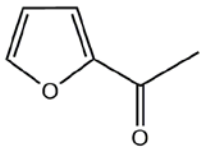
Acetic acid, **34**



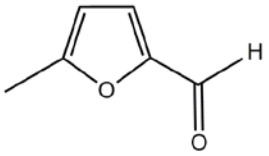
3,5-Diethyl-2-methylpyrazine, **39**



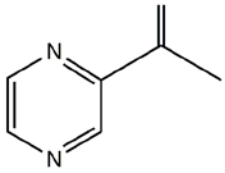
Furfuryl formate, **41**



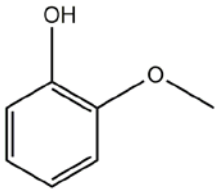
Acetylfuran, **42**



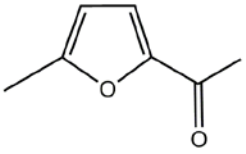
5-Methyl furfural, **49**



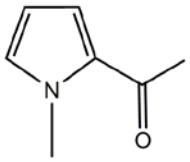
1-Methylethenylpyrazine, **51**



Guaiacol, **82**



2-Acetyl-5-methylfuran, **83**



1H-Pyrrole-2-carboxaldehyde, **91**

Table of Contents Graphic