

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/5632148>

# When Machine Tastes Coffee: Instrumental Approach To Predict the Sensory Profile of Espresso Coffee

ARTICLE *in* ANALYTICAL CHEMISTRY · APRIL 2008

Impact Factor: 5.64 · DOI: 10.1021/ac702196z · Source: PubMed

CITATIONS

46

READS

164

7 AUTHORS, INCLUDING:



**David Labbe**

Nestlé S.A.

20 PUBLICATIONS 350 CITATIONS

SEE PROFILE



**Andreas Rytz**

Nestlé S.A.

32 PUBLICATIONS 750 CITATIONS

SEE PROFILE



**Chahan Yeretdzian**

Zurich University of Applied Sciences

88 PUBLICATIONS 1,973 CITATIONS

SEE PROFILE



**Imre Blank**

Nestec

122 PUBLICATIONS 3,227 CITATIONS

SEE PROFILE

# When Machine Tastes Coffee: Instrumental Approach To Predict the Sensory Profile of Espresso Coffee

Christian Lindinger,<sup>\*,†</sup> David Labbe,<sup>†</sup> Philippe Pollien,<sup>†</sup> Andreas Rytz,<sup>†</sup> Marcel A. Juillerat,<sup>†</sup> Chahan Yeretian,<sup>†,‡</sup> and Imre Blank<sup>†,§</sup>

Nestlé Research Center, Vers-Chez-les-Blanc, 1000 Lausanne 26, Switzerland

A robust and reproducible model was developed to predict the sensory profile of espresso coffee from instrumental headspace data. The model is derived from 11 different espresso coffees and validated using 8 additional espressos. The input of the model consists of (i) sensory profiles from a trained panel and (ii) on-line proton-transfer reaction mass spectrometry (PTR-MS) data. The experimental PTR-MS conditions were designed to simulate those for the sensory evaluation. Sixteen characteristic ion traces in the headspace were quantified by PTR-MS, requiring only 2 min of headspace measurement per espresso. The correlation is based on a knowledge-based standardization and normalization of both datasets that selectively extracts differences in the quality of samples, while reducing the impact of variations on the overall intensity of coffees. This work represents a significant progress in terms of correlation of sensory with instrumental results exemplified on coffee.

The perception elicited from drinking a freshly prepared espresso coffee represents a complex scientific phenomenon.<sup>1,2</sup> This multisensory experience involves all our senses such as olfaction, taste, texture, trigeminal, and visual sensation. Furthermore emotions and cognitive processes constructed during drinking experiences, such as interactions between senses<sup>3</sup> and product familiarity,<sup>4</sup> modulate perception. Among the various sensory modalities, the aroma (smell) and taste, often referred to as flavor, are of paramount importance to the quality of coffee. The flavor compounds in a roast and ground (R&G) coffee depend on many factors, two of which are of particular importance. First, the green coffee variety and quality with its specific composition on precursors sets the stage for the later flavor development during roasting. Second, the roasting process which unlocks the

flavor potential of the green coffee beans and creates the coffee flavor so much appreciated by coffee aficionados all over the world. Changes in these two factors affect most of the flavor compounds.

Trained sensory panelists are capable to characterize subtle differences between espresso coffees. The set of sensory descriptors that are employed by a coffee sensory specialist are generally adjusted to the type of coffee being evaluated and pertain either to green coffee tasting, the two commercial coffee varieties being Robusta (species *Coffea canephora*) and Arabica (species *Coffea arabica*), or to finished product tasting such a specific espresso blend from a coffee roaster.

Besides sensory analysis, coffee scientists have long been searching for instrumental approaches to complement and eventually replace human sensory profiling. Yet, the prediction of sensory profiles on the basis of instrumental data (e.g., PTR-MS) has remained a challenge that still waits to be resolved.<sup>5–8</sup> Most attempts to relate aroma perception to analytical measurements probably failed because one was trying to establish a too direct relationship between the two datasets that are fundamentally different in nature. Indeed, (i) a sensory profile reflects relations among a set of attributes measured on a relative scale with respect to references, while the choice of attributes and references is to some degree arbitrary. It is primarily influenced by differences in the sensory qualities of samples (i.e., the relative composition) rather than from the absolute intensity of physical stimuli (i.e., absolute quantities/concentrations). (ii) In contrast, analytical measurements (e.g., ion trace profiles in PTR-MS) mainly reflect information about absolute concentrations of individual constituents, essentially independent of variations from other constituents (PTR-MS ion signal) or from other samples.

In order to correlate sensory and instrumental datasets, it is important to apply a knowledge-based standardization and normalization procedure to both datasets that selectively filters out mutually relevant information.

Considerable work has been devoted in the last 20 years to the characterization and quantification of aroma-active compounds

\* Author to whom correspondence should be addressed. Phone: +41 21 785 9384. E-mail: christian.lindinger@rdls.nestle.com.

<sup>†</sup> Nestlé Research Center, Lausanne, Switzerland.

<sup>‡</sup> Present address: Nestlé Nespresso SA, Paudex, Switzerland.

<sup>§</sup> Present address: Nestlé Product Technology Center, Orbe, Switzerland.

(1) Digum, G.; Luttinger, N. *The Coffee Book. Anatomy of an Industry: From the Crop to the Last Drop*; The New Press: New York, 1999.

(2) Pendergrast, M. *Uncommon Grounds. The History of Coffee and How It Transformed Our World*; Basic Books: New York, 1999.

(3) Murphy, C.; Cain, W. S.; Bartoshuk, L. M. *Sens. Processes* **1977**, 1 (3), 204–211.

(4) Koster, E.P. *Chem. Sens.* **2005**, 30 (Suppl 1), i236–i237.

(5) Lindinger, W.; Hansel, A.; Jordan, A. *Int. J. Mass Spectrom. Ion Processes* **1998**, 173, 191–241.

(6) Pollien, P.; Jordan, A.; Lindinger, W.; Yeretian, C. *Int. J. Mass Spectrom.* **2003**, 228, 69–80.

(7) Lindinger, W.; Hirber, J.; Paretzke, H. *Int. J. Mass Spectrom. Ion Processes* **1993**, 129, 79–88.

(8) Hansel, A.; Jordan, A.; Holzinger, R.; Prazeller, P.; Vogel, W.; Lindinger, C. *Int. J. Mass Spectrom. Ion Processes* **1995**, 149/150, 609–619.

in coffee as well as to their interactions. In parallel, many sensory profiling methods have been developed to describe the rich aroma profile of coffee and to identify sensory relevant compounds by chemical analysis.<sup>6</sup> Roasted coffee aroma is composed of more than 1000 volatile organic compounds (VOC).<sup>9</sup> Since the application of gas chromatography coupled with olfactometry (GC-O) as a screening tool in aroma analysis, it is known that only about 50 volatile compounds contribute to the overall sensory mapping of coffee aroma and even less need to be taken into account to characterize and reproduce quite accurately a coffee aroma.<sup>10</sup> These are present in a subtle ratio to provide a balanced coffee aroma. Consequently, most of the VOC in the headspace of coffee have probably no aroma contribution or a threshold too high to be perceived by humans.<sup>11–13</sup>

To measure the headspace concentration of these compounds on-line with a high reproducibility, a mass spectrometer is needed with low fragmentation, high time resolution, and broad linear detection range. Among others a PTR-MS-type gas analyzer fulfils these demands. Furthermore the development of a “hot liquid headspace inlet system” was necessary to analyze the aroma compounds released from espresso coffee (headspace) in a reproducible way. The coupling of a commercial GC/MS (electron impact, EI, ionization) with PTR-MS (chemical ionization) finally allows for achieving an unambiguous identification and builds the bridge between molecular information and on-line marker based headspace data, applicable to complex food systems.<sup>14</sup>

This study is aiming to give an answer to a fundamental question of flavor science: “Can we predict the sensory profiles of coffee by a fast (high-throughput) on-line analytical measurement?” Or more colloquially “Can a machine taste coffee?”

We demonstrate that this is indeed possible, provided that both analytical and sensory datasets are properly standardized and normalized with respect to each other, in order to selectively extract the mutually relevant “quality” information.

## TECHNICAL SECTION

A standard espresso machine (TURMIX C250, Zurich, Switzerland) was used for the preparation of espresso coffee. Extraction time, quantity, and temperature were kept stable for all instrumental measurements and the sensory evaluation. For the development of the model 11 espresso coffees (“Ristretto”-type) were used (training set). Subsequently the model was tested on a set of eight different espresso coffees (validation set) where five of them were “Ristretto”-types and three of them “Lungo”-types. The coffees were extracted with 25 mL of water (“Ristretto”-type) and with 110 mL (“Lungo”-type) of water (Vittel) as these are the recommended extraction volumes for these espressos. The extraction time was limited to  $20 \pm 2$  s for the “Ristretto”-type

and  $40 \pm 4$  s for the “Lungo”-type. Samples exceeding these limits were excluded from further analysis.

**Instrumental Analysis.** A double-jacketed, water-heated sample cell (350 mL glass vessel) was mounted inside an oven set to 65 °C with active air circulation (Figure 1). A temperature stabilized water bath (set to 50 °C) was connected to the double-jacketed cell to keep the sample at constant temperature. The sample cell was connected to the fix-mounted top cover to be easily disconnected and filled with the coffee sample. The coffee headspace was purged continuously with 300 standard cubic centimeters per minutes (sccm) through heated tubes (80 °C) penetrating the fix-mounted top of the cell. Before analyzing by PTR-MS, the sample gas was diluted with 3000 sccm dry air preventing saturation of the instrument. Lindinger et al.<sup>14</sup> have described the complete setup in detail. After preparation of the coffee sample, the headspace vessel was disconnected, filled with the coffee sample, and connected back to the system. This procedure was carried out quickly (less than 10 s) to avoid temperature changes of the sample, oven, and headspace vessel. PTR-MS instrumental parameters were set as following: drift tube pressure, 2 mbar; drift tube temperature, 80 °C; drift voltage, 550 V; extraction blend, 6 V.

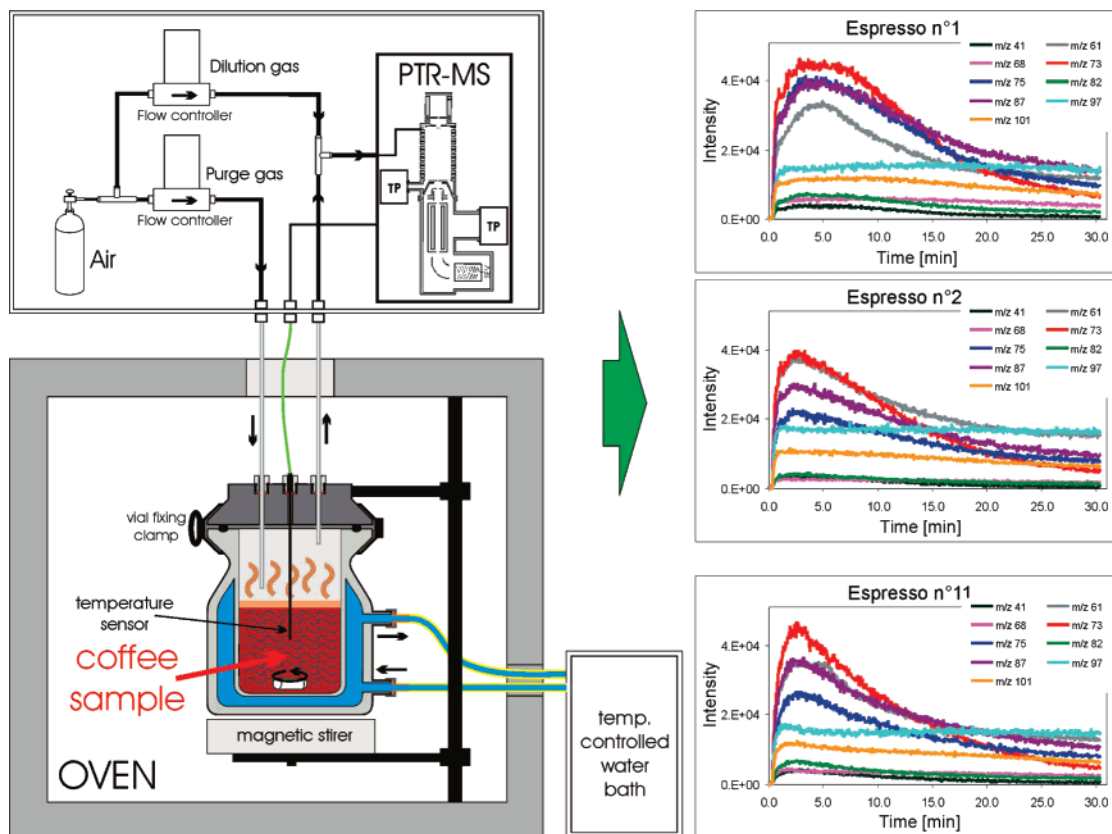
As we were aiming at a high-throughput experimental solution, it was important to first reduce the instrumental data to include only the most relevant information. This concerns the measurement time period per sample as well as the specific ion traces to be included.

First, scans of the ion trace intensities over the mass range from  $m/z$  20 to  $m/z$  250 were performed for 30 min after connecting the sample to the analytical setup. The maximum ion trace intensities in the headspace were measured 2 min after connecting the sample cell to the PTR-MS. From there on, and for all subsequent measurements, only these maximum concentration values were used for the data treatment, and the total measurement time was reduced to 3 min.

Second, statistical data analysis of all coffee samples measured with three repetitions allowed selecting 16 most discriminating ion traces out of 230, as listed in Table 1. This was based on the  $F$  ratio of the one way analysis of variance (ANOVA). This reduced data set of 16 ion traces was used for further statistical analyses and correlation. This reduction of the analytical data to a limited, most significant fraction was crucial to achieve a reproducible and robust correlation with the sensory data.

It is also important to point out that the chemical identity of the 16 ion traces is not relevant for this study, and in particular the correlation is not based on a set of identified key aroma compounds. Most of the odor active compounds in coffee are indeed known and can be analyzed and quantified with modern instrumental techniques. Yet, the aim of this work was to demonstrate the applicability of a data-driven method rather than a targeted chemical study. Furthermore, by applying a fast on-line technique, we were aiming at developing a high throughput solution. It is interesting that some of the 16 ion traces selected in our holistic approach do indeed, at least partially, represent well-known coffee aroma compounds, such as isobutanal, 3- and 2-methylbutanal, 2,3-butanedione, and 2,3-pentanedione, for example.

- (9) Nijssen, L. M.; Visscher, C. A.; Maarse, H.; Willemsen, L. C.; Boelens, M. H. *Volatile Compounds in Food*; TNO Nutrition and Food Research Institute: Zeist, The Netherlands, 1996.
- (10) Pollen, P.; Krebs, Y.; Chaintreau, A. *Conference Proceedings: ASIC-17eme Colloque Scientifique International sur le Café*; 1997.
- (11) Czerny, M.; Mayer, F.; Grosch, W. *J. Agric. Food Chem.* **1999**, *47*, 695–699.
- (12) Mayer, F.; Czerny, M.; Grosch, W. *Eur. Food Res. Technol.* **2000**, *211*, 272–276.
- (13) Mayer, F.; Grosch, W. *FFJ* **2001**, *16*, 180–190.
- (14) Lindinger, C.; Pollen, P.; Ali, S.; Yeretdzian, C.; Märk, T. *Anal. Chem.* **2005**, *77*, 4117–4124.



**Figure 1.** Schematic presentation of the experimental setup with on-line sampling of coffee headspace using a dynamic headspace cell. Examples of the obtained time intensity profiles for three different espressos are shown on the right side.

However, the combination of off-line coupled GC/MS/PTR-MS and on-line PTR-MS permit us to achieve both a targeted chemical analysis as well as a high throughput study.<sup>14</sup> The ionization process in PTR-MS (chemical ionization) is relatively soft. Nevertheless, 90% of the identified molecules suffer from some fragmentation, and each of the 16 recorded ion traces represents the sum of more than one chemical compound (including fragments) at different concentrations. Table 1 shows the list of compounds representing 95% of the ion trace signal. As an example, the ion trace at  $m/z$  57 represents more than 95% acetol while  $m/z$  81 represents the presence of pyridine, pyrazine, 2-furfuryl formate, 2-furfuryl acetate, and 2-furfuryl alcohol in a relative abundance.

**Sensory Analysis.** The sensory evaluation was performed by an experienced 10-member coffee panel. Even though the panel was already trained with the glossary and products, four training sessions were organized to select the appropriate sensory attributes. During training, two or three references were presented for each attribute (weak, medium, and high intensity). Tasters were asked to rank and score the intensity level. Prior to the final evaluation in sensory booths, the performance of the panel and consensus between individual tasters was checked.

Two performance indices were used, i.e., discrimination ability of the panel and individual performance of assessors. The discrimination ability was tested for each attribute, using a two-way ANOVA with the product as fixed and the assessor as a random factor. The  $F$ -ratio of the product effect served as marker of the discrimination ability. The individual performance of an assessor on a single attribute consisted of the individual discrimi-

nation ability and the ability to be consensual with the other assessors of the panel.<sup>15</sup>

The individual discrimination ability was tested using a one-way ANOVA with product as a fixed factor. The assessor was considered as able to discriminate products whenever the  $F$ -ratio was larger than 2.5 (8 products evaluated twice). Discrimination ability also implies repeatability of the assessor. The consensus ability was tested by calculating the correlation coefficient between individual product means and the product means of the rest of the panel. An assessor was judged consensual whenever this correlation coefficient was larger than 0.6 (8 products,  $p$ -value = 0.1).

The Quantitative Descriptive Analysis<sup>16</sup> method with the monadic approach was applied for performing the sensory evaluation of the coffees, i.e., samples were presented one by one. Each assessor scored the product based on the knowledge and consensus acquired during the training sessions. The samples were presented using a balanced presentation design, which was repeated twice. The products were evaluated in booths using a scale ranging from 0 (not intense) to 10 (very intense).

**Mathematical Data Treatment.** Standard statistical methods (ANOVA) were used to describe the discrimination of the products by sensory analysis. From this, the quality of the predicting model based on instrumental data was evaluated. In addition, a validation using different samples was performed

(15) Labbe, D.; Rytz, A.; Morgeneegg, C.; Ali, S.; Martin, N. *Chem. Sens.* **2007**, 32, 205–214.

(16) Stone, H.; Sidel, J. L. *Sensory Evaluation Practices*; Academic Press: London, U.K., 1993.

**Table 1. List of Compounds Representing 95% of the Individual Ion Trace Signals, the Compounds Were Identified by GC/MS**

Volatile compound	PTR-MS ion trace [m/z]														
	41	45	57	61	68	69	73	75	81	82	83	87	89	97	101
acetaldehyde															
methyl formate															
propanal															
furan															
isobutanol															
methyl acetate															
methylpropanoate															
2-methylfuran															
3-methylfuran															
2-butanone															
2-methylbutanal															
3-methylbutanal															
2,5-dimethylfuran															
2,3-butanedione															
2,3-pentanedione															
hexanal															
N-methylpyrrole															
pyridine															
pyrazine															
2-methyltetrahydrofuranone															
3-hydroxy-2-butanone															
acetol															
acetic acide															
2-furfural															
2-furfuryl formate															
2-acetyl-furan															
pyrrole															
1-(2-furyl)-2-propanone															
2-furfuryl acetate															
5-methyl-2-furfural															
γ-butyrolactone															
2-furfurylalcohol															

and tested by comparing the LSD obtained from the sensory evaluation (validation set) and the difference between the predicted and evaluated data for each sensory attribute and each product. The predata-processing of the raw data is described in Table 2A,B. Standard matrix transformations are applied. Principal component regression (PCR) was applied to obtain the matrix transforming the instrumental data into predicted sensory data. The approach is described in detail using standard mathematics NCSS 2007 (Hintze, J. *NCSS, PASS, and GESS*; NCSS: Kaysville, Utah, 2006) and Excel (Microsoft) was used to do the calculations presented in this work.

## DEVELOPING THE SENSORY PREDICTIVE MODEL

In a first part of the study, 11 espresso coffees (training set) of distinct sensory profiles were analyzed both by sensory and analytical methods, in order to develop the correlation model. Subsequently, in a second part, the model was tested on a set of 8 different coffees (validation set).

**Sensory Evaluation.** Ten trained panelists evaluated each coffee twice by applying Quantitative Descriptive Analysis.<sup>17</sup> Nine aroma and two taste attributes were scored on a scale ranging from 0 (not intense) to 10 (very intense) for each product by monadic sensory evaluation. The statistical significance of each sensory attribute was validated by processing an analysis of variance (ANOVA). As the performance of each panelist was confirmed prior to the sensory evaluation session, attribute means (mean of all evaluations, including the repetitions, for one attribute and one coffee) are used for further data treatment, Table 2.  $x_{pu}$

represents the matrix elements of the sensory raw data matrix  $X$  for 11 coffees ( $p$ ) and 11 sensory attributes ( $u$ ). For standardization of the sensory evaluation data between 0 and 1, the following transformation was applied to  $X$ , Table 2B

$$X'_{pu} = (x_{pu} - x_{u \min}) / (x_{u \max} - x_{u \min}) \quad (1)$$

where  $x_{pu}$  = mean value (all panelists) of the  $u$ th attribute of coffee  $p$ ,  $x_{u \max} = \max_{i=1, \dots, 11} (x_{iu})$ ,  $x_{u \min} = \min_{i=1, \dots, 11} (x_{iu})$ .

The resulting matrix  $X'$  describes primarily the quality of the different coffees. In addition, it also includes information about the intensity of each sample. Furthermore, it was observed that the range of scale covered by panelists during the evaluation changes when presenting very different coffees in terms of intensity. Finally, it should be considered that the quality of the coffees may strongly influence the evaluation of the differences in overall intensity. A panelist very sensitive to bitterness will evaluate a very bitter coffee as very intense too. Therefore, attention has to be paid to the effect of diluting the same samples (same quality but different intensity) and/or including coffees with very different overall intensities (different quality and different intensity). The same coffee diluted with water should show the same quality by sensory evaluation but the intensities should be lower.

In order to make sensory data more amendable to correlation with instrumental data, information other than that on quality differences should be selectively eliminated/reduced (the same holds for instrumental data).

In psychophysics, there is ample empirical evidence that the relationship between the magnitude of a sensory sensation and

(17) Stone, H.; Sidel, J.; Oliver, S.; Woosley, A.; Singleton, R. C. *Food Technol.* **1974**, *28*, 24–34.



**Table 2. Standardization and Normalization of the Instrumental and Sensory Data Including the Correlation of Both Datasets, Only a Fraction of the Full Dataset Is Shown**

**A: Instrumental data**

M Raw data	m/z 41	m/z 45	m/z 57	m/z 61	m/z 68	m/z 69	m/z 73	m/z 75	m/z 81	m/z 82	m/z 83	m/z 87	m/z 89	m/z 97	m/z 101	m/z 111
Espresso n°1	144.9	3392	73.6	750.0	165.0	746.5	1374	785.7	1021	254.2	375.5	1081	57.6	311.1	284.8	239.6
Espresso n°2	127.4	3074	70.1	855.3	107.7	639.9	1237	647.3	780.7	170.7	298.9	935.9	49.2	328.0	283.8	229.4
⋮																
Espresso n°11	116.1	3150	69.0	834.3	109.9	563.9	1110	619.5	819.6	169.1	251.9	831.4	48.9	350.6	289.4	245.5
column mean:	130.6	3238	71.1	894.4	111.2	676.4	1253	701.8	866.4	177.4	312.3	950.7	53.4	384.3	293.8	260.1
<div><math>m'_{pk} = m_{pk} / \text{Mean}_{k=1..11} (m_k)</math></div>																
M'																
Espresso n°1	1.109	1.048	1.036	0.839	1.484	1.104	1.097	1.120	1.178	1.433	1.203	1.137	1.080	0.810	0.969	0.921
Espresso n°2	0.976	0.949	0.986	0.956	0.969	0.946	0.987	0.922	0.901	0.962	0.957	0.984	0.921	0.854	0.966	0.882
⋮																
Espresso n°11	0.889	0.973	0.971	0.933	0.988	0.834	0.886	0.883	0.946	0.953	0.807	0.874	0.915	0.912	0.985	0.944
<div><math>m''_{pk} = m'_{pk} / c_p</math> where <math>c_p = \text{Mean}_{k=1..16} (m'_{pk})</math></div>																
M''																
Espresso n°1	1.010	0.954	0.943	0.764	1.352	1.005	0.999	1.020	1.073	1.306	1.095	1.035	0.984	0.737	0.883	0.839
Espresso n°2	1.032	1.005	1.043	1.012	1.026	1.001	1.045	0.976	0.954	1.018	1.013	1.042	0.975	0.903	1.022	0.934
⋮																
Espresso n°11	0.968	1.059	1.058	1.016	1.076	0.908	0.964	0.961	1.030	1.038	0.878	0.952	0.997	0.994	1.073	1.028
Normalization																
M'''																
Espresso n°1	0.154	-0.431	-0.887	-1.457	1.390	0.083	0.007	0.124	0.604	1.669	0.775	0.684	-0.322	-1.085	-1.080	-1.313
Espresso n°2	0.408	0.043	0.581	0.058	0.088	0.047	0.785	-0.129	-0.444	0.090	0.136	0.804	-0.486	-0.393	0.191	-0.547
⋮																
Espresso n°11	-0.338	0.558	0.793	0.081	0.290	-0.767	-0.573	-0.215	0.227	0.198	-0.906	-0.861	-0.079	-0.016	0.650	0.218

**B: Sensory data**

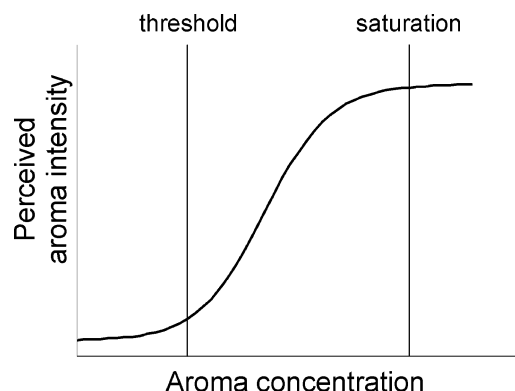
X Raw data	Coffee	Bitter	Cocoa	Roasted	Woody	Cereal	ButterToffee	Acid	Citrus	Winey	Flowery
Espresso n°1	5.79	7.35	2.08	5.47	2.34	2.17	1.57	2.99	1.47	0.72	0.15
Espresso n°2	5.28	3.38	0.89	4.98	1.92	1.38	2.78	4.26	2.04	1.41	1.05
⋮											
Espresso n°11	5.22	4.64	1.30	4.88	2.44	2.13	2.17	4.29	2.20	1.61	0.98
Xmax:	6.06	7.96	2.65	6.00	2.94	2.84	3.04	6.16	4.94	3.44	3.89
Xmin:	4.27	3.07	0.46	3.73	0.81	1.26	0.85	2.61	0.68	0.62	0.13
X'	$x'_{pu} = (x_{pu} - x_{u \min}) / (x_{u \max} - x_{u \min})$										
Espresso n°1	0.85	0.88	0.74	0.76	0.72	0.58	0.33	0.11	0.18	0.03	0.01
Espresso n°2	0.57	0.06	0.20	0.55	0.52	0.07	0.88	0.46	0.32	0.28	0.24
⋮											
Espresso n°11	0.53	0.32	0.38	0.50	0.76	0.55	0.60	0.47	0.36	0.35	0.23
X''	$x''_{pu} = x'_{pu} / c_p \text{ where } c_p = \text{Mean}_{k=1...16} (m'_{pk})$										
Espresso n°1	0.78	0.80	0.67	0.70	0.65	0.52	0.30	0.10	0.17	0.03	0.00
Espresso n°2	0.60	0.07	0.21	0.58	0.55	0.08	0.93	0.49	0.34	0.29	0.26
⋮											
Espresso n°11	0.58	0.35	0.42	0.55	0.83	0.60	0.66	0.52	0.39	0.38	0.25
X'''	Normalization										
Espresso n°1	0.90	1.27	0.95	0.74	0.36	0.55	-1.18	-0.97	-0.76	-0.98	-0.93
Espresso n°2	0.30	-0.94	-0.69	0.40	0.02	-0.99	0.86	0.21	-0.07	-0.11	-0.12
⋮											
Espresso n°11	0.23	-0.08	0.05	0.30	0.97	0.79	-0.02	0.28	0.14	0.17	-0.16

**C: Correlation of the two datasets**

Step 1: PCA on analytical data	$M'''_{(11,16)}$ = Transformed analytical data (normalized). $M'''_{(11,16)} = P_{(11,10)} * L_M^T_{(10,16)}$ (PCA=single value decomposition). P is the matrix of scores. $L_M$ is the matrix of loadings.
Step 2: Generalized inverse	$P_{(11,10)} = M'''_{(11,16)} * L_M_{(16,10)} * (L_M^T_{(10,16)} * L_M_{(16,10)})^{-1}$
Step 3: Principal component regression (PCR)	$X'''_{(11,11)}$ = Transformed sensory data (normalized). $X'''_{(11,11)} = P_{(11,10)} * L_X^T_{(10,11)}$ In this step, the matrix P is known from step 1 and the matrix $X'''$ is directly derived from the sensory data. The matrix $L_X$ is obtained by using the generalized inverse of P: $L_X^T_{(10,11)} = (P^T_{(10,11)} * P_{(11,10)})^{-1} * P^T_{(10,11)} * X'''_{(11,11)}$
Step 4: Full model by combining results of steps 2 and 3	$X'''_{(11,11)} = P_{(11,10)} * X'''^T_{(10,11)}$ $X'''_{(11,11)} = M'''_{(11,16)} * L_M_{(16,10)} * (L_M^T_{(10,16)} * L_M_{(16,10)})^{-1} * L_X^T_{(10,11)}$
Step 5: Model selection	Since there are fewer samples than masses, the system is saturated. If the model has to be used for prediction purpose, this full model is clearly over-fitting the data. It is therefore necessary to select a model that does not over-fit the data. A family of models is given by using the hierarchy of principal components (i.e. include only the first m principal components, where m is a number between 1 and 10). $X'''^m_{(11,11)} = M'''_{(11,16)} * L_M_{(16,m)} * (L_M^T_{(m,16)} * L_M_{(16,m)})^{-1} * L_X^T_{(m,11)}$

the intensity of the physical stimulus is best represented by a sigmoid curve. This relationship is expected to be equally valid for the perceived aroma intensity vs aroma concentration. The

sigmoid relationship is characterized by three parts, Figure 2: (i) The perceived intensity is zero for concentrations below the perception threshold. (ii) The perceived intensity is almost



**Figure 2.** Schematic presentation of the sigmoid relationship between perceived aroma intensity and aroma concentration (Fechner, 1877).

constant above a certain (high) concentration. The perception is at saturation, and higher concentration does not lead to stronger perception above the saturation point. Variation in the quality of the perceived sensory note is yet possible with increasing concentration above saturation. (iii) The perceived intensity increases with concentration between these two limits, without changes in the quality of the perceived sensory note.

For the intermediate part, Fechner (1877) law states that the intensity of a sensation increases as the logarithm of the stimulus:

$$\text{perception} \sim k \log(\text{concentration}) \quad \text{“Fechner’s law”}$$

This basic law was adapted by many psychophysicists. Stevens,<sup>18</sup> for example, proposed a power law, but the general concept of the sigmoid relationship is well accepted.

This concept is nevertheless only valid when evaluating a single stimulus at different concentrations (e.g., a coffee extracted at 25 mL, with or without addition of water). When analyzing different

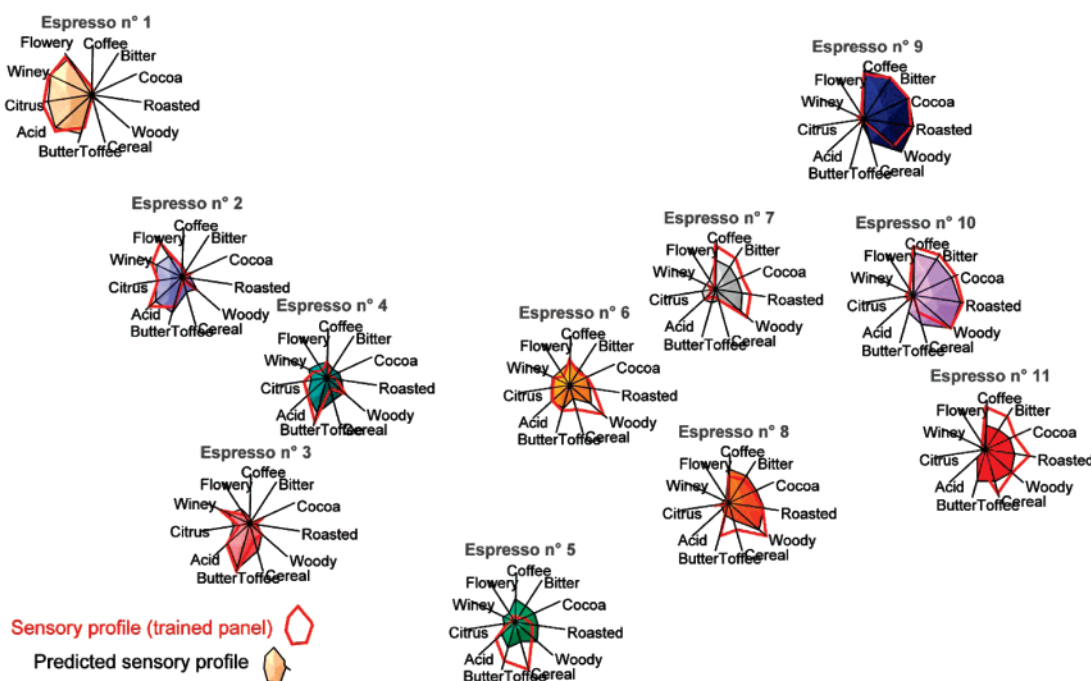
stimuli (different coffees), the quality of perception is changing and it is no longer possible to evaluate their intensity on a single (same) scale.

The challenge for extending this basic concept of psychophysics is twofold: (i) find a way to rate the perception of different coffees on a common scale system (refers to sensory) and (ii) separate composition and concentration (refers to sensory and analytical).

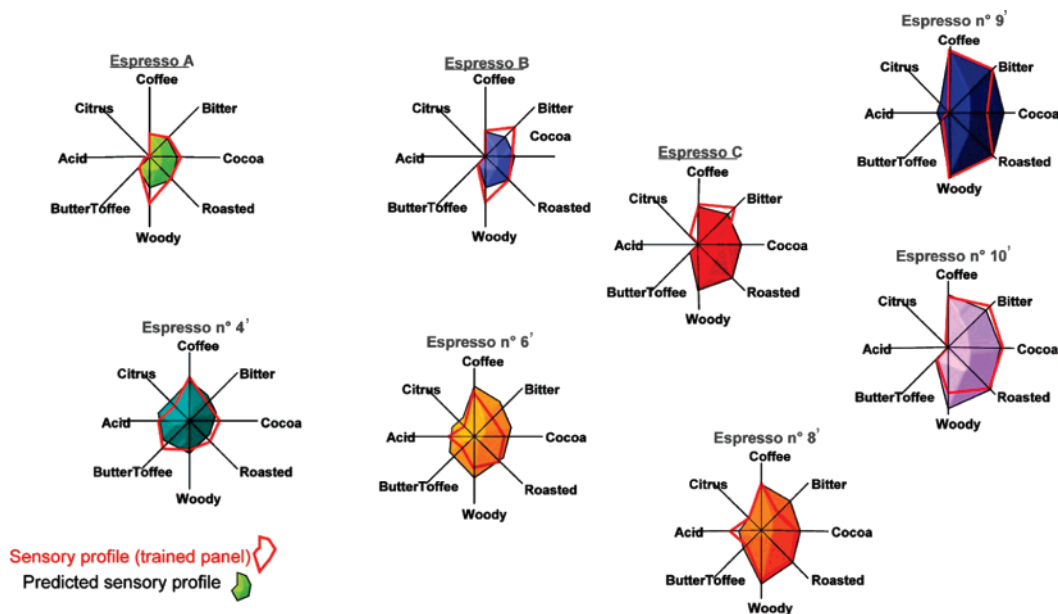
Therefore, a multivariate scale system was applied for sensory evaluation to compare mean profiles (mean from repetitions of a product by one panelist, is considered here as the raw sensory data) for (i) any coffee with any other coffee (quality) and (ii) any concentration of a coffee with any other concentration of the same coffee (quantity).

Referring to the matrix  $\mathbf{X}'$ , it is impossible to separate quality data from differences in quantity using this dataset. A sensory panelist is able to evaluate the intensity of different coffees if the coffee product remains the same and, for example, only the dilution is changed. As soon as the coffee is very different in quality it is hardly possible to evaluate its intensity consistently because there are individual differences as to which sensory attribute will be the reason for the perceived intensity. One panelist may evaluate a bitter coffee as intense, while another panelist might be more influenced by the acidity. We have therefore to conclude that the overall intensity (quantity) of a product can hardly be separated from the profile (quality) in sensory profiling.

Here we have to refer to the instrumental data set which contains quantity/ intensity information. As an approximation, we take the integrated intensity over a PTR-MS ion trace spectrum for a given product, as a measure of the overall intensity (quantity) of the specific espresso. Consequently, both instrumental and sensory data are standardized to the quantity information calculated from the instrumental data. This is one key element that



**Figure 3.** Results of the model development showing the determined sensory profiles and predicted sensory profiles. The data was standardized to the minimum and maximum score of the sensory evaluation for each attribute ( $\mathbf{X}'$ ).



**Figure 4.** Validation of the model using eight coffees ( espressos A, B, C, nos. 4', 6', 8', 9', and 10'). Espresso nos. 4', 6', 8', 9', and 10' are the same as used for the training set but harvested 2 years later. The data was standardized to the minimum and maximum score of the sensory evaluation for each attribute ( $\mathbf{X}'$ ).

allows correlating sensory quality with instrumental quality data, once the impact of the overall intensity of the products has been filtered out in a consistent way from both data sets.

**Instrumental Considerations.** The headspace of 11 different espresso coffees (training set) was monitored on-line by PTR-MS. The intensity maxima in headspace concentrations were reached 2 min after connecting the temperature-controlled sample cell to the analytic setup. At this point 16 ion traces were quantified for each coffee (see Table 1).

Similar to the sensory profiles where each attribute is evaluated by using the same range of scale, the instrumental data have to be standardized before extracting the information on quantity differences.  $m_{pk}$  represents the matrix elements of instrumental raw data matrix  $\mathbf{M}$  for 11 coffees ( $p$ ) and 16 ion traces ( $k$ ). If the analytical data consisted of a quantification of individual chemical compounds, their individual sensory thresholds could be used for standardizing the data. With the obtained ion-trace profiles, no such information is available since each ion-trace may represent more than one compound (Table 1) and the compound's contribution individually changes among the coffees. Thus, standardizing to the sum of the 16 ion-traces ( $\text{sum}_{i=1...16}(m_{pi})$ ) would result in a relative overestimation of very intense aroma compounds related to a specific ion trace. On the contrary, relatively low concentrated aroma compounds would contribute to a negligible extent. Therefore, it is proposed to give the same weight to each ion-trace, independently of its absolute abundance. One way to achieve this is standardizing all ion trace intensities to the mean value across the different coffees as shown in Table 2A:

$$m'_{pk} = m_{pk} / \text{mean}_{i=1...11}(m_{ik}) \quad (2)$$

where  $m_{ik}$  is the value of the  $k$ th ion trace of coffee  $i$ .

Further mathematical transformation of the ratios between abundances of a given ion trace has the advantage that it is in-line with Fechner's law.

On the basis of the standardized matrix  $\mathbf{M}'$ , the overall intensity (quantity) of each sample can be expressed. A more intense coffee will generally contain higher intensities for all recorded ion-traces. Therefore, the data can either be standardized by using the sum of all ion-trace intensities of a single coffee or by using the mean value. Tests showed better results in correlation when using the mean value of all ion-trace intensities resulting in a  $\mathbf{C}$  vector with

$$c_p = \text{mean}_{k=1...16}(m'_{pk}) \quad (3)$$

where  $m'_{pk}$  is the value of the  $k$ th standardized ion trace of coffee  $p$ .

**Linking Sensory and Instrumental Data.** With the use of the  $\mathbf{C}$  vector (eq 3), both the instrumental and sensory data can be transformed to contain only information reflecting the quality differences between the coffees.

$$m''_{pk} = m'_{pk} / c_p \quad (4)$$

where  $\mathbf{M}'$  is the standardized instrumental data set matrix with its elements  $m'_{pk}$

$$x''_{pu} = x'_{pu} / c_p \quad (5)$$

where  $x'_{pu}$  is the standardized sensory data set.

With normalization of the transformed sensory and analytical data sets, they can now be correlated ( $\mathbf{M}''$  and  $\mathbf{X}''$  from Table 2). The matrix of scores is obtained by applying a PCA (principal components analysis) to the transformed analytical data. To predict the sensory profile based on the analytical data, the obtained matrix of scores of the transformed analytical data replaces the

(18) Stevens, S. S. *Psychol. Rev.* **1957**, *64*, 153–81.



matrix of scores of the PCA applied to the transformed sensory data. Table 2C shows the complete calculation, step by step. In general, a fairly good prediction of the sensory profile was achieved based on analytical data, i.e., the general sensory profile is well reflected by the model using the first two principal components (Figure 3). As an example, the sensory and instrumental data of espressos no. 1 and no. 10 resulted in an almost complete overlap, reflecting well the sensory difference between these coffees described as flowery, winey, citrus, and acid and coffee, bitter, cocoa, roasted, and woody, respectively. In some cases, the intensity of certain sensory attributes is deviating from those obtained by sensory evaluation (e.g., espresso no. 11); however, the trends are still well represented by the model. The accuracy of the model can be determined by comparing the difference between predicted and evaluated sensory data. It was observed that for each sensory attribute and each coffee, the difference is smaller than the LSD (least-squares deviation) obtained from the sensory evaluation data (applying ANOVA).

After having established the predictive model, on the basis of 11 espresso coffees, it was validated on a set of 8 additional coffees.

### VALIDATING THE MODEL

Eight coffees were used for the validation of the predictive model. Five of them were the same as used in the training set, but the beans were harvested 2 years later. Three were new espresso coffees, extracted with 110 mL of water ("Lungo"-type), in contrast to all other espresso samples that were extracted with 25 mL of water ("Ristretto"-type). The application of the model to the validation set resulted in an overall median absolute difference between observed and predicted of 0.11 on the 0–1 scale, which is much smaller than the LSD of all attributes. The sensory evaluation of the validation samples was performed using a reduced set of eight sensory attributes because of the missing significance of the three remaining attributes (Figure 4). Again, there is a good correlation between the sensory profiles with those obtained by instrumental analyses. The espresso nos. 4', 6', 8', 9', and 10' showed profiles similar to those harvested 2 years earlier (espresso nos. 4, 6, 8, 9, 10, Figure 3), even though a reduced set of sensory attributes were used. Moreover, the sensory profiles of the new coffees (espresso A–C) could also be predicted based on PTR-MS data, thus validating the model, despite the fact that these samples were prepared at higher dilution ("Lungo"-type). These results indicate that the predictive model is robust enough to withstand variations derived from harvest, instrumental changes, and preparation/extraction.

### CONCLUSIONS

The coffee extraction and all measurements were performed under well-defined conditions to achieve high reproducibility of data. A requirement of this study was to perform coffee extraction, sensorial mapping, and analytical measurements under conditions that are close to the real consumption event. Nevertheless, there is a difference between the sensory evaluation of trained panelists and the description given by an average coffee consumer.<sup>19,20</sup> For this study it was important to generate a complete evaluation of the sensory experience rather than a description using a reduced

set of attributes such as overall aroma, roasty aroma, processy aroma, body, persistency, etc.

The standardization procedure described in the results chapter (eq 1) is not the only possibility to generate a common sensory evaluation data set using a number of individual sensory evaluations. Another possibility would be to normalize the individual evaluation of each panelist and attribute before performing the same standardization. A panelist who is not able to describe the difference for a specific attribute within the presented coffees might evaluate all samples around the same point within the scale. By normalization of such a dataset, it would have the same importance as a dataset of a panelist well discriminating the samples using the specific attribute; thus the result would be biased. In other words, if panelists are using the range of scale in a very different way, e.g., one using only the upper part but evaluating the differences in the right order and another one using the complete range of scale, a mean value of both might generate incoherent results. Though there are several approaches to data normalization, it turned out that using the attribute means as shown in this work allows for obtaining the most reproducible results.

Performing correlations between headspace data (volatiles only) and sensory data implies the challenge of modeling taste attributes such as bitter and acid. However, because of the fact that during the roasting process both volatiles and nonvolatiles are formed by similar reaction pathways, it is not surprising to obtain correlations with taste attributes represented by nonvolatiles on the basis of a volatiles pattern. Therefore, taste attributes can indeed be predicted by using data of volatiles only.<sup>21,22</sup>

The espresso coffees selected for this study were blended mainly from different Arabica varieties with some Robusta (5% Robusta in average). Blending Arabica with Robusta varieties has shown to add complexity to the model and adds some uncertainty in the prediction of the sensory profiles, but the model still shows reliable results. Models obtained with pure Robusta and Arabica samples separately (data not shown here) show results which are more accurate. Nevertheless, predicting the sensory profile of a new coffee depends very much on the diversity of samples used for constructing the model. The broader the sensory and analytical space of the samples included in developing the model, the more robust the model. We can further deduce that the sensory attributes citrus, flowery, acid, bitter, and winey are best discriminating the products and at the same time they were predicted most accurately by the instrumental data.

In conclusion, this study has resulted in a first robust model to predict sensory profiles of coffee from analytical data. Furthermore, as the model is based on fast on-line PTR-MS analysis, a prediction of a sensory profile can be accomplished within minutes. Relative to current methods of aroma profile analysis, this opens the possibility of high throughput studies.

Received for review October 24, 2007. Accepted November 23, 2007.

AC702196Z

(21) Frank, O.; Zehentbauer, G.; Hofmann, T. *Eur. Food Res. Technol.* **2006**, *222*, 492–508.

(22) Frank, O.; Blumberg, S.; Kunert, C.; Zehentbauer, G.; Hofmann, T. *J. Agric. Food Chem.* **2007**, *55*, 1945–1954.

(19) Roberts, A. K.; Vickers, Z. M. *J. Sens. Stud.* **1994**, *9*, 1–2.

(20) Moskowitz, H. R. *J. Sens. Stud.* **1996**, *11*, 19–37.