



# Integrated $^1\text{H}$ NMR fingerprint with NIR spectroscopy, sensory properties, and quality parameters in a multi-block data analysis using ComDim to evaluate coffee blends

Michel Rocha Baqueta<sup>a</sup>, Aline Coqueiro<sup>a,b</sup>, Paulo Henrique Março<sup>a</sup>, Manuela Mandrone<sup>c</sup>, Ferruccio Poli<sup>c</sup>, Patrícia Valderrama<sup>a,\*</sup>

<sup>a</sup> Universidade Tecnológica Federal do Paraná, Campus Campo Mourão (UTFPR-CM), Campo Mourão, Paraná, Brazil

<sup>b</sup> Universidade Tecnológica Federal do Paraná, Campus Ponta Grossa (UTFPR-PG), Ponta Grossa, Paraná, Brazil

<sup>c</sup> University of Bologna, Department of Pharmacy and Biotechnology (FaBiT), Bologna, Italy

## ARTICLE INFO

### Keywords:

Coffee quality  
ComDim-based method  
Simultaneous data analysis  
 $^1\text{H}$  NMR-based metabolomics  
Industrial case study  
Portable NIR

## ABSTRACT

Coffee quality is determined by several factors and, in the chemometric domain, the multi-block data analysis methods are valuable to study multiple information describing the same samples. In this industrial study, the Common Dimension (ComDim) multi-block method was applied to evaluate metabolite fingerprints, near-infrared spectra, sensory properties, and quality parameters of coffee blends of different cup and roasting profiles and to search relationships between these multiple data blocks. Data fusion-based Principal Component Analysis was not effective in exploiting multiple data blocks like ComDim. However, when a multi-block was applied to explore the data sets, it was possible to demonstrate relationships between the methods and techniques investigated and the importance of each block or criterion involved in the industrial quality control of coffee. Coffee blends were distinguished based on their qualities and metabolite composition. Blends with high cup quality and lower roasting degrees were generally differentiated from those with opposite characteristics.

## 1. Introduction

The advances in the analytical instrument areas and the beginning of chemometric tools for extracting information from complex data sets have reached all areas of science and are also present and in continuous progress in the field of food science and technology. Concerning the analytical evaluation of foods, coffee is a complex chemical matrix that deserves extensive studies and explorations. In industrial quality control, coffee is evaluated by several factors, including chemical, physical and sensory determinations (Baqueta, Coqueiro, Março, & Valderrama, 2020; Baqueta, Coqueiro, & Valderrama, 2019; Craig, Botelho, Oliveira, & Franca, 2018; Ribeiro, Ferreira, & Salva, 2011). Although the combination of these criteria is often used to guarantee the quality and identity of coffee, the importance of these various criteria and their relationships has not been investigated in the industrial context.

The chemical composition of coffee dictates its flavor and aroma, meaning its cup profile, and consequently its price during trading and export (Farah, Monteiro, Calado, Franca, & Trugo, 2006; Franca, Mendonça, & Oliveira, 2005; Ribeiro et al., 2011). In Brazil, for example,

*Coffea arabica* L. (Arabica) is classified by cup quality in different categories, including hard, rioysh, and rio, while *Coffea canephora* (var. Robusta) is classified from excellent to abnormal (Brazil, 2003). However, in the country, many companies combine beans of different cup qualities to produce their commercial coffees, which increases the complexity of the coffee composition for analytical purposes. Blends are not only made to create new flavor profiles and meet demand according to consumer preferences but also to standardize quality and adjust the final coffee price (Craig et al., 2018; Dias et al., 2018).

In the literature, some analytical approaches have been proposed to evaluate coffee blends, either to determine and predict their chemical composition (Assis, Pereira, Amador, Augusti, de Oliveira, & Sena, 2019; Jumhawan, Putri, Yusianto, & T., & Fukusaki, E., 2016; Souard, Delporte, Stoffelen, Thévenot, Noret, Dauvergne, & Stévigny, 2018) or to identify and quantify adulterations (Bertone, Venturello, Giraud, Pellegrino, & Geobaldo, 2016; Combes, Joët, & Lashermes, 2018; Wermelinger, D'Ambrosio, Klopprogge, & Yeretian, 2011). However, to date, no study has been carried out to characterize the quality and composition of coffee blends with different cup and roasting profiles, a

\* Corresponding author at: Universidade Tecnológica Federal do Paraná (UTFPR), 87301-899, Campo Mourão, Paraná, Brazil.

E-mail address: [patriciav@utfpr.edu.br](mailto:patriciav@utfpr.edu.br) (P. Valderrama).

subject that deserves attention.

Currently, the demand of the modern industry has propagated several claims and arguments to develop fast, simple, robust, and eco-friendly analytical methods with industrial application in the coffee market (Assis et al., 2019; Baqueta et al., 2020, 2019; Barbin, Felício, Sun, Nixdorf, & Hirooka, 2014; Correia et al., 2018; Craig et al., 2018; dos Santos, Alvarenga, & Boffo, 2020; Milani, Rossini, Catelani, Pezza, Toci, & Pezza, 2020). These solutions include instrumental techniques such as proton Nuclear Magnetic Resonance ( $^1\text{H}$  NMR) spectroscopy and Near-Infrared (NIR) spectroscopy (including portable and bench-top NIR equipment), which have achieved popularity for solving problems related to the coffee analytical evaluation (Barbin et al., 2014; Consonni, Polla, & Cagliani, 2018; Correia et al., 2018; Monteiro et al., 2018; Toci et al., 2018).  $^1\text{H}$  NMR and NIR spectroscopy are complementary techniques, where  $^1\text{H}$  NMR provides metabolite fingerprints and complements the molecular information obtained with NIR spectroscopy. However, NIR equipment is relatively cheaper, simpler, automated, and less “information-heavy” than  $^1\text{H}$  NMR, especially when it is a portable NIR, making it more accessible for laboratory studies and routine analysis.

Multiple product measurements, as in assessing coffee quality, have become easier today and, in the chemometric domain, multi-block data analysis methods are valuable for integrating and studying multiple information describing the same samples, extracting interesting features of multiple data blocks, in an exploratory way (Cariou, Jouan-Rimbaud Bouveresse, Qannari, & Rutledge, 2019; El Ghaziri, Cariou, Rutledge, & Qannari, 2016; Mishra et al., 2020). Common Component and Specific Weights Analysis (CCSWA), whose algorithm is currently called ComDim (for Common Dimension), is an exploratory tool that belongs to the family of multi-block methods (Bouveresse, Pinto, Schmidtke, Locquet, & Rutledge, 2011; Cariou et al., 2019; El Ghaziri et al., 2016; Qannari, Wakeling, Courcoux, & MacFie, 2000). Its main idea is to replace several separate multivariate exploratory analyses in each matrix (e.g. principal component analysis - PCA), or a data fusion followed by PCA, with a single multi-block data analysis applied to all of them, with the advantage to bring the ‘salience’ of each block indicating the importance of each one and its variability in each dimension (Bouveresse et al., 2011; Mishra et al., 2020). In the literature, ComDim has been applied to characterize complex samples in several ways (Bouveresse et al., 2011; Loudiyi, Rutledge, & Ait-Kaddour, 2018; Martin et al., 2015; Ríos-Reina, Callejón, Savorani, Amigo, & Cocchi, 2019; Rosa et al., 2017). However, the potential of multi-block methods is little explored due to the lack of experience of many analysts (Mishra et al., 2020). Therefore, further investigations need to be done to explore the capacity of ComDim, especially in the context of the quality of coffee that would be interesting for industrial application.

In view of the main factors involved in the coffee quality control in an industrial context, this case study hypothesized that a ComDim exploratory analysis could integrate and evaluate metabolite fingerprints (obtained with  $^1\text{H}$  NMR), NIR spectra, sensory properties, and quality parameters of coffee blends of different cup and roasting profiles, to explore their qualities and composition. Moreover, using the ComDim multi-block method, it was tried to establish relationships between multiple data blocks and understand the importance of each technique and method of analysis used in the coffee industry. Finally, the ability of the ComDim method was compared to a multivariate exploratory analysis using data fusion-based PCA.

## 2. Materials and methods

### 2.1. Sampling and industrial processing

The samples of this study were provided by a coffee producer cooperative from Paraná State (Brazil). The coffee cherries were from the 2016/2017 and 2017/2018 crops and were obtained by dry process. The green beans were mechanically dried (CIA Lilla, Brazil) to a

moisture content of 12% and dehusked using a coffee hulling machine (CIA Lilla, Brazil). Further details on the green coffee processing by this cooperative can be found in the literature (Baqueta et al., 2017). Pre-processed green coffees showed known differences in cup quality based on the Brazilian classification (Brazil, 2003). The samples of green Arabica coffee were previously classified by the cup as hard, rioysh, and rio, while Robusta was classified as excellent. A simplified description of how the raw beans were processed is provided below.

In the industry, four blends (from different batches of around 225 kg each) were routinely manufactured with different roasting profiles and based on the previous classification of green coffees. The beans were roasted in an industrial Lilla roaster, model OPUS 40 (CIA Lilla, Brazil). The hard/rioyssh blend with a medium-dark roasting degree was roasted at 240 °C for 15 min (coffee blend 1), hard/rioyssh blend with a dark roasting degree was roasted at 260 °C for 15 min (coffee blend 2), rioyssh/rio blend with a medium-dark roasting degree was roasted at 255 °C for 15 min (coffee blend 3) and hard/Robusta blend with a light roasting degree was roasted at 210 °C for 20 min (coffee blend 4). The roasted beans were then ground in an industrial Lilla grinding roller, model Orion (CIA Lilla, Brazil) with fine adjustment.

During sample collection, performed at random to include inter- and intra-batch variations in manufactured coffee blends, the same process conditions mentioned above were followed. A total of forty-eight genuine samples (0.5 kg per sample) were collected and considered in this study, representing the industrial process where tons of coffee were being produced. From these, fifteen samples were collected for coffee blend 1, fifteen were from coffee blend 2, and fifteen were from coffee blend 3. From coffee blend 4, which is a superior coffee, three samples were collected due to the lowest production in the sector. All samples have the purity and quality seals of the Brazilian Coffee Industry Association (ABIC). Coffee blend 1 has the traditional seal, coffee blends 2 and 3 have the extra strong seal, and coffee blend 4 has the premium seal.

It is worth mentioning that coffee blends are real samples, manufactured over several days and represent the inherent complexity and real variations in coffee processing. In addition, the analyses made on the samples are described below. All measurements (physicals, chemical, and sensorial) were made with a single analysis, and realized in the routine analysis of the coffee industry.

### 2.2. Traditional analyses in the coffee industry

Four quality parameters (granulometry, color, infusion time, and moisture content) were evaluated in all coffee blend samples ( $n = 48$ ) as is usually done in the industry, using methods described previously in the literature for the same samples (Baqueta et al., 2020). Briefly, granulometry was measured using 100 g of each sample on a vibrating platform (Bertel, Caieiras, Brazil) for 10 min and 2 mm in amplitude. Color measurement was performed on a Minolta colorimeter (Chroma Meter CR-410, Japan) with standard  $D_{65}$  lighting and normal colorimetric viewing angle. The colorimeter was calibrated to provide results similar to those of the Agtron Roast Color Classification System. The color was expressed in parameters developed by the *Commission Internationale de l'Éclairage* (CIE): luminosity ( $L^*$ ), green-red component ( $a^*$  coordinate), and blue-yellow component ( $b^*$  coordinate). To evaluate the infusion time, coffee powders (100 g) were extracted with filtered water (1000 mL) at 92 °C by percolation with Melitta 103 filter papers (Guaíba, Brazil). The infusion time parameter, which means the coffee extraction time, was measured from the beginning to the end of the beverage extraction. The moisture content was determined in 3 g of roasted and ground coffee samples using an Infrared Moisture Analyzer (Gehaka IV2000) in “auto-dry” mode at 105 °C.

The professional cupping was also applied as is usually done in the Brazilian industry, using a method previously described for the same samples (Baqueta et al., 2019) and according to technical recommendations/standards from ABIC (ABIC, 2018). Due to the participation of

cuppers, this work was authorized by the Research Ethics Committee of UTFPR – Paraná – Brazil, under protocol number 2.810.398.

Each coffee beverage ( $n = 48$ ) was prepared for the cupping procedure: the coffee powders (100 g) were extracted with filtered water (1000 mL) at 92 °C by percolation with Melitta 103 filter papers (Guafba, Brazil). The beverages were evaluated by two professional cuppers who work commercially at the coffee cooperative and have more than 20 years of experience in coffee cupping. The professional cuppers were men from 50 to 55 years old, hired by the coffee industry to evaluate coffee beverages. The cuppers tasted coffee beverages two to four times and evaluated nine specific sensory properties, including powder fragrance, drink aroma, acidity, bitterness, flavor, body, astringency, residual flavor, and overall quality (ABIC, 2018). The final notes were given using a scale presented in our previous study (Baqueta et al., 2019), where 0 represents an “unpleasant coffee” and 10 an “excellent coffee”.

### 2.3. Spectral techniques

In addition to the traditional methods described above, the coffee blend samples were investigated by two instrumental techniques: portable NIR spectroscopy and a  $^1\text{H}$  NMR-based metabolomics approach. These techniques were included to complement the characterization of the samples, obtaining more specific information on the metabolite composition of coffees with  $^1\text{H}$  NMR and exploring the potential of portable NIR to monitor coffee quality in a faster and more accessible way.

#### 2.3.1. Portable NIR spectroscopy

To avoid the need for previous pretreatments, coffee powders ( $n = 48$ ) contained in the packages were analyzed directly using a portable NIR spectrometer (microNIR™ 1700) from JDSU Uniphase Corporation with a glass cuvette in reflectance mode. The reflectance spectra of each sample were recorded from 906 to 1676 nm (32 scans, step of 6.20 nm). The blank was evaluated using a standard NIR reflectance (Spectralon™) with a diffuse reflection coefficient of 99%, while a dark reference (zero-to simulate non-reflection) was obtained with the lamp off. The NIR spectra were recorded from an aliquot of approximately 1 g of each sample of ground coffee.

#### 2.3.2. $^1\text{H}$ NMR-based metabolomics

The coffee samples ( $n = 48$ ) were analyzed using a  $^1\text{H}$  NMR-based metabolomic approach. For this procedure, all reagents were purchased from Sigma Aldrich, except deuterated solvents ( $\text{H}_2\text{O}-d_2$  and 3-(trimethylsilyl)-propionic-2,2,3,3- $d_4$  acid sodium salt–TMSP purchased from Eurisotop. The aqueous extracts were obtained by simulating the conditions used to prepare the coffee beverage: one hundred milligrams of the ground and roasted coffee were extracted using 1.5 mL of phosphate buffer (90 mM, pH 6.0) in  $\text{H}_2\text{O}-d_2$  containing 0.01% of TMSP as standard, for 1 h in a preheated bath maintained at a fixed temperature of 90 °C.  $^1\text{H}$  NMR spectra were recorded at 298 K using a Varian 14.4 T NMR instrument (600.13 MHz operating at  $^1\text{H}$  frequency) equipped with a high-field triple resonance probe, using  $\text{H}_2\text{O}-d_2$  for the internal lock. Relaxation delay of 2.0 s observed pulse of 5.80  $\mu\text{s}$ , and the sum of 256 scans were acquired for each sample. The acquisition time was 16 min and the spectral width of 16.00 ppm. A presaturation sequence was used to suppress the residual water signal at 4.83 ppm (power = 22 Hz, presaturation delay = 2 s).

### 2.4. Preliminary data processing and first evaluation

Before performing multi-block data analysis, information on quality parameters and sensory properties were evaluated by simple descriptive statistics, such as average and range, using Microsoft Excel 2013, to give an idea of the variation of these results for each coffee blend.

For  $^1\text{H}$  NMR spectra, free induction decays (FIDs) were Fourier

transformed, and the resulting spectra were phased, baseline-corrected, and calibrated for TMS at 0.00 ppm. The spectral intensities were reduced to integrated regions of equal width (0.04 ppm) corresponding to the region of 0.00 to 12.00 ppm with scaling on the standard at 0.00 ppm using the NMR MestReNova software (Mestrelab Research, Spain). The regions from 5.00 to 4.50 ppm were excluded from the analysis due to residual water signals. Binning was performed using MestReNova software, normalized to the standard and each bin was the average sum of 0.04 ppm intervals. First, the  $^1\text{H}$  NMR data were analyzed visually to identify differences between the spectra of the four coffee blends and to identify their metabolites. The identification of metabolites in  $^1\text{H}$  NMR fingerprints was based on the chemical shifts, coupling constants, and comparison with data from the available literature on coffee (Toci et al., 2018; Wei, Furihata, Miyakawa, & Tanokura, 2014).

For portable NIR spectroscopy, the reflectance spectra were transformed into absorbance and pre-processed by multiplicative scatter correction (Geladi, MacDougall, & Martens, 1985) and Savitzky–Golay smoothing (Savitzky & Golay, 1964) with a 5-point window and first-order polynomial. This data pre-processing performed through Matlab software version R2007b (The MathWorks Inc., Natick, USA) was used.

Additionally, the data sets were concatenated in a matrix (with the same pre-processing described below in Section 2.5. *Multi-block data analysis*) to identify possible outliers and assess the overall quality of the data. Possible outliers were assessed using a test of leverage against Q residuals. While leverage represents how far a sample is from the data center, Q residuals represent non-modeled residuals. The samples can certainly be considered outliers when they simultaneously show high leverage and high Q residuals (Dias et al., 2018). On the other hand, PCA was carried out to explore the concatenated matrix. More information about the PCA can be found in the literature (Ferreira, 2015).

### 2.5. Multi-block data analysis

The main idea of this work is to use ComDim multi-block data analysis for the simultaneous study of multiple sets of matrices with different variables that describe the coffee blend samples. In multi-block methods, multivariate data transformation techniques are specific to highlight the contributions of different data blocks (Mishra et al., 2020). The ComDim procedure begins by organizing the data into blocks, where each block contains data related to a technique or method used, resulting in four data blocks investigated in this study. The first block contained the information of the  $^1\text{H}$  NMR spectra (chemical shifts), the second of the NIR spectra (absorbance), the third of the quality parameters (granulometry, infusion time, moisture content, and  $L^*$ ,  $a^*$ , and  $b^*$ ), and the fourth of the sensory property notes (powder fragrance, drink aroma, acidity, bitterness, flavor, body, astringency, residual flavor, and overall quality) determined in each sample.

The quality parameters were auto-scaled, while the sensory properties, NIR, and  $^1\text{H}$  NMR spectra were normalized by the standard deviation. After concatenation, each block was normalized by its Frobenius' norm so that they all have the same total variance (Bouveresse et al., 2011; El Ghaziri et al., 2016; Mishra et al., 2020).

The ComDim algorithm used in this study was developed and coded for the Matlab software and the SAISR Toolbox can be downloaded for free at <http://www.chimiometrie.fr/saisirdownload.html> (Cordella & Bertrand, 2014). Details about the algorithm are presented point by point in the literature (Bouveresse et al., 2011; Cariou et al., 2019; El Ghaziri et al., 2016; Qannari et al., 2000). In general, for many multivariate data analysis techniques, as in PCA, a weighted sum of the ‘variable variance-covariance’ matrix is usually calculated. In the ComDim-based method, however, the main intention is to calculate a weighted sum of the ‘sample variance-covariance’ matrix of each block, and then extract its first principal component standardized as the first ‘Common Dimension’ (CD) or ‘Common Component’ (CC). The algorithm extracts the global and local components from multiple blocks of data sequentially and then iteratively calculates the weight or ‘salience’

of each block for the calculated CC. 'Salience' indicates the importance of each block in the construction of the CC, and it is then possible to calculate a percentage of the variability extracted by each dimension or CC. After calculating the first CC, each original data block matrix is deflated, and the procedure is repeated for calculating the second CC, and so on. Thus, each CC is the first PC of a weighted sum of the 'sample variance-covariance' of deflated matrices (Bouveresse et al., 2011; Cariou et al., 2019; El Ghaziri et al., 2016; Qannari et al., 2000). The ComDim analysis provides informative graphs showing similarities and differences between the samples through scores, the importance of each block and its relationships through saliences, and the origins of variables responsible for the similarities and differences observed in the samples through loadings. As well as the clustering of samples reveals similarities between the samples and the clustering of variables indicates inter-variable relationships, in the ComDim method, similar salience proportions indicate inter-block relationships (Bouveresse et al., 2011; Cariou et al., 2019; El Ghaziri et al., 2016; Rosa et al., 2017).

### 3. Results and discussion

#### 3.1. Quality parameters evaluation

The individual results of the four quality parameters determined in each sample are available in **Table S1** as **Supplementary Material** and a simple descriptive statistic, including average and range for each coffee blend, are presented in **Table S2**. In general, the samples presented granulometry from 0.60 to 0.85 mm (from 80.95% to 93.27% in 100 g of sample), infusion time ranged from 3.28 to 5.23 min, and moisture content from 3.10 to 5.50%. For the roasting degree, evaluated by luminosity parameter ( $L^*$ ), the samples of coffee blend 1 had an average of 51.80, while samples of blend 2 had an average of 42.20. The samples of coffee blend 3 had an average luminosity of 46.01, and blend 4 an average of 58.64. This shows that the coffee blends have a well-defined roasting degree considering the average values and generally follow a pattern. This was expected because the industry uses different roasting profiles for each coffee blend, but maintains a pattern of grinding, moisture, and infusion time. The average results for the four quality parameters are within the identity and quality standard of the industry that supplied the samples (Baqueta et al., 2020) and following ABIC technical recommendations for commercial coffees (ABIC, 2018).

#### 3.2. Coffee beverage evaluation

Regarding professional cupping, the result of each sensory property for each sample is presented in **Table S3** in the **Supplementary Material**. The simple descriptive statistic, including the average and range of each coffee blend, are presented in **Table S4**. The final notes differed extensively for coffee blends 1, 2, and 3 compared to blend 4. Coffee blend 4 received the highest notes for all analyzed attributes with overall quality equal to 7.00, followed by coffee blend 1 (6.05), 2 (5.81), and 3 (5.04). The sensory quality of coffee blends is within the technical recommendations of ABIC for different categories (ABIC, 2018): traditional or extra strong coffees must have an overall quality from 4.5 to 5.9 (coffee blends 1, 2, and 3) and premium from 6.0 to 7.2 (coffee blend 4).

**Fig. S1** annexed on **Supplementary Material** shows the average results of specific sensory properties for each coffee blend and its profiles. Most of the properties analyzed received very close notes to each other, which justified the overlapping data. Also, it was observed that the notes attributed to the overall quality of the coffee blends were related to the initial profile of the cup and the degrees of roasting. Blend 3 (rioysh / rio medium-dark) with low cup quality and blend 2 (hard-rioysh dark roast) with the high roasting grade received lower notes, while blends 1 and 4 with better cup quality and lighter roasting degrees (mixtures 1 and 4) received higher notes. This indicates or suggests that the final note of coffee blends is highly dependent on the cup profile of the raw coffee before blending and the roasting conditions to which they are subjected.

However, lower cup quality does not mean absolutely that coffee is not of quality, but that it has lower points in professional cupping. Besides, this often does not reflect the preference of coffee consumers (Giacalone, Fosgaard, Steen, & Münchow, 2016).

#### 3.3. Qualitative evaluation of $^1\text{H}$ NMR and NIR spectra

Before performing the chemometric analysis, the spectral data were examined visually. **Fig. 1** shows a representative  $^1\text{H}$  NMR fingerprint for each studied coffee blend ranging from 0.00 to 9.50 ppm. **Fig. 2A and B** show the  $^1\text{H}$  NMR spectra of a representative sample with attribution of the identified metabolites. **Fig. S2** showed in the **Supplementary Material** shows the  $^1\text{H}$  NMR fingerprints of all samples plotted with Matlab software and used in the ComDim analysis.

In general,  $^1\text{H}$  NMR fingerprints showed a complex resonance pattern, since there are many different compounds present in the coffee blend extracts. At the first glance, the four  $^1\text{H}$  NMR fingerprints looked extremely similar. Based on previous investigations of coffee metabolites (Toci et al., 2018; Wei et al., 2014) and detailed analyses of chemical shifts and patterns, it was possible to attribute the major metabolites present in the coffee blends. The signals at  $\delta$  3.29, 3.49, 3.89, and 7.85 were attributed to caffeine; at  $\delta$  4.41, 8.04, 8.57, and 8.81 to *N*-methylpyridinium; at  $\delta$  4.45, 8.12, 8.84, and 9.13 to trigonelline; at  $\delta$  8.49 to formate; at  $\delta$  1.33, and 4.15 to lactate; at  $\delta$  1.96 to acetate; at  $\delta$  3.33, 3.65, and 4.09 to *myo*-inositol; at  $\delta$  1.88, 1.99, 2.15, 3.55, 4.03, and 4.14 to quinic acids; at  $\delta$  2.53 to  $\gamma$ -quinide; at  $\delta$  3.21 to choline; and the regions at  $\delta$  0.82–0.99 to lipids; and  $\delta$  6.28–7.66 to chlorogenic acids (CGAs). However, as it is not possible to determine qualitative differences in the metabolite content of coffee blends by visual inspection, the  $^1\text{H}$  NMR fingerprints were subsequently explored by ComDim analysis.

**Fig. 3** shows the NIR spectra of all samples. Observing the changes in the NIR spectra, it can be seen that all samples showed absorption bands at all wavelengths; however, a higher absorption at the third (900–1000 nm), second (1100–1200 nm), and first (1400–1500 nm) overtones can be observed. Although some molecular vibrations can be attributed to the NIR spectra, the complex composition of coffee makes it difficult to indicate which compounds are responsible for the absorption bands (Barbin et al., 2014). In general, absorption bands can be attributed to most components found in coffee, such as caffeine, lipids, sucrose, and other carbohydrates, proteins, trigonelline, chlorogenic acids, and phenols (Barbin et al., 2014; Ribeiro et al., 2011). The absorption bands around 900–1000 nm can be attributed to lipid vibrations (Barbin et al., 2014) and also to phenolic compounds (Correia et al., 2018; Monteiro et al., 2018), which are products of the thermal process degradation of CGAs (Farah et al., 2006; Franca et al., 2005). The absorptions on 1100–1200 nm can be attributed to vibrations of caffeine, lipids, and carbohydrates (Barbin et al., 2014). Other carbohydrate vibrations can also be observed in the region from 1400 to 1500 nm (Barbin et al., 2014). For the region from 1400 to 1676 nm, the absorption bands can be attributed to chlorogenic acids, phenols, alcohols, water, and lipids in coffee (Barbin et al., 2014; Correia et al., 2018; Ribeiro et al., 2011). However, as the changes along the NIR spectra are not visible, the application of chemometrics is again necessary to determine possible differences between coffee blend profiles.

#### 3.4. Outlier identification and data fusion evaluation by PCA

Possible outliers were investigated among the samples. According to **Fig. S3A** presented in the **Supplementary Material**, no sample presented simultaneously high leverage and high Q residuals, indicating that potential outliers were not identified in the samples. Also in **Supplementary Material**, **Fig. S3B** shows PCA scores plot obtained in a concatenated matrix using the same pre-processing data applied in the multi-block data analysis. The first two principal components (PCs) explained the major variation observed in the data structure (>98%, 96.12% in PC1, and 2.02% in PC2). Based on the PCA scores distribution



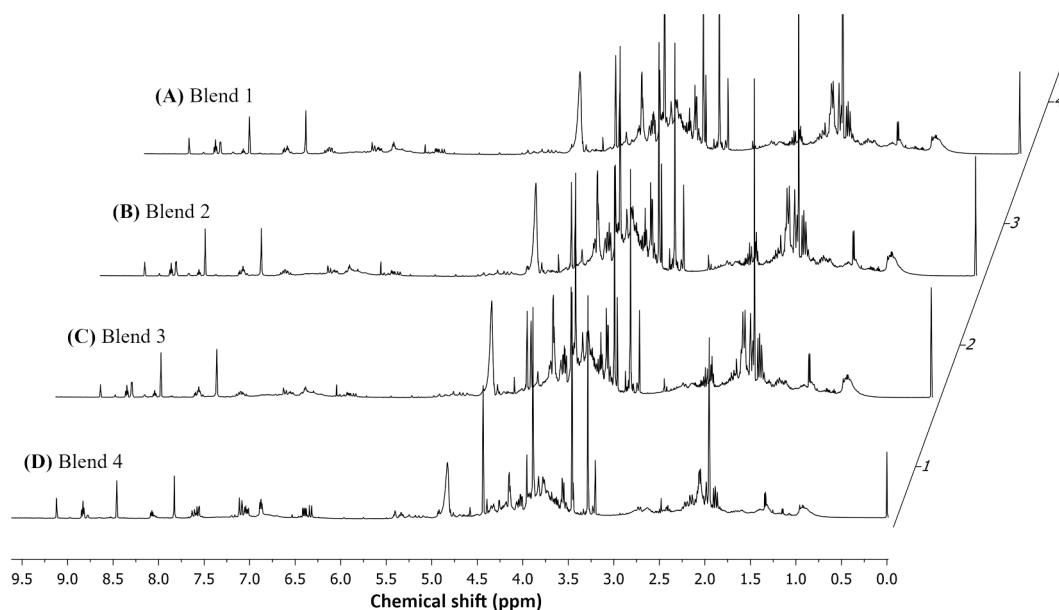


Fig. 1.  $^1\text{H}$  NMR fingerprints. Coffee blend 1, hard/rioysh medium-dark roast (A); Coffee blend 2, hard/rioysh dark roast (B); Coffee blend 3, rioysh/rio medium-dark roast (C); and Coffee blend 4, hard/Robusta light roast (D).

in PC1, there is no clear differentiation of the coffee blends because all samples presented positive scores. On the other hand, PCA scores in PC2 demonstrated that all samples from coffee blend 4 presented positive scores and were distinguished from other coffee samples. There is also some separation between the samples of blend 1 that generally presented positive scores, while the samples of coffee blends 2 and 3 presented generally negative scores. However, the PCA does not make it possible to explore the importance of each data block to characterize the samples and explore their inter-block relationships, which makes it unfavorable compared to the ComDim, which performs data analysis simultaneously, providing the importance of each data set in a multi-block study.

### 3.5. ComDim-based method

The main idea of this study was to verify the ability of the ComDim analysis to explore the multiple data blocks that describe coffee blends and understand the importance of each technique/method, and its relationships in the industrial context. For this,  $^1\text{H}$  NMR fingerprints (block 1), NIR spectra (block 2), quality parameters (block 3), and professional cupping results (block 4), were considered.

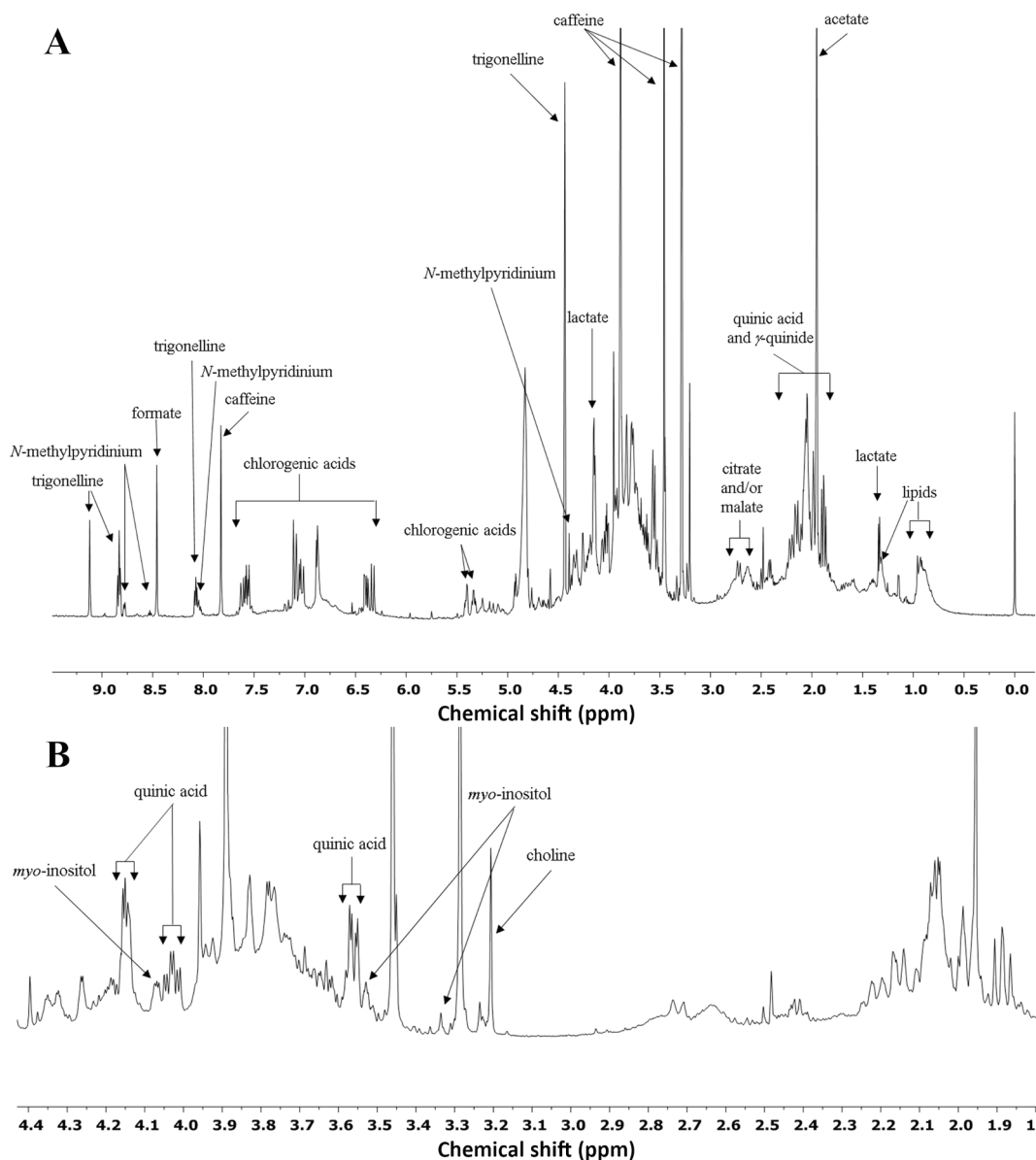
Fig. 4 shows the results obtained with the ComDim analysis, where the informative graphics show the saliences (Fig. 4A), scores (Fig. 4B), and loadings (Fig. 4C). In the ComDim procedure, the scores, saliences, and loadings are calculated for each CC or common dimension. Five CCs were computed on this ComDim analysis and considered sufficient to take into account all sources of variation in the four data blocks (Fig. 4A). No significant information was observed in more than five CCs. From CC2 to CC5 a little variability was observed, which reinforces the no need for more CCs in the model. In general, in common dimensions from CC2 to CC4, the greatest salience was only for quality parameters (block 3) while in CC5 the greatest salience was only for  $^1\text{H}$  NMR spectroscopy. As the main idea is to verify relationships among the data blocks (i.e. similar salience proportions indicating inter-block relationships), the results of these CCs were not interesting because it does not show similar saliences (importance) in the 4 blocks simultaneously. The quality parameters were of great importance in CC2, CC3, and CC4, probably associated with the degree of roasting (color parameters –  $L^*$ ,  $a^*$ , and  $b^*$ ) well-defined for each coffee blend, as previously mentioned.

In the exploratory analysis ComDim, the first CC (CC1) showed the

searched results, demonstrating natural clusters of samples, the importance of each block, and the relationships between the information obtained from the portable NIR, the  $^1\text{H}$  NMR fingerprints, the sensory properties and the quality parameters of coffee blends. With a ComDim analysis applied to the blocks, it was possible to demonstrate the importance of each criterion involved in the industrial quality control of coffee and to distinguish profiles of coffee blends and their characteristics. Therefore, only the scores and loadings referring to the projection in CC1 were discussed below.

For CC1 (Fig. 4A), saliences show that all blocks of data contribute to the dispersion behavior of the samples, showing importance in this decreasing order: portable NIR (block 2), sensory properties (block 4),  $^1\text{H}$  NMR spectroscopy (block 1), and quality parameters (block 3). Portable NIR spectroscopy was the main block to distinguish the qualities of coffee blends, as it showed the greatest salience; however, the sensory properties, the fingerprints of the metabolites obtained with  $^1\text{H}$  NMR, and the quality parameters also showed importance, indicating that the information contained in the blocks are complementary and can have an interesting synergistic effect. The quality parameters (block 3) showed less contribution compared to the other blocks, showing that only the parameters color ( $L^*$ ,  $a^*$ , and  $b^*$ ) and moisture content were relevant in the ComDim analysis, while granulometry and infusion time were not significantly related. This suggests that granulometry and infusion time parameters are not very important to differentiate the coffee blends. However, it is possible to establish a relationship between the coffee roasting degree and the information obtained by other techniques and methods.

The scores of ComDim (Fig. 4B) were evaluated to highlight possible clusters according to their characteristics and it was observed that the coffee blend samples were differentiated based on their qualities and metabolic composition. Coffee blends with high cup quality and low degrees of roasting (blend 1 – samples from 1 to 15, and blend 4 – samples from 46 to 48) were generally differentiated from those with opposite characteristics: lower cup quality and a high degree of roasting (blend 2 – samples from 16 to 30, and blend 3 – samples from 31 to 45). Samples of coffee blend 1 and 4 had generally negative scores, while samples of coffee blend 2 and 3 had more positive scores. Furthermore, these clusters were more consistent than when compared to the PCA scores, since it is possible to see a clearer marginal separation between the samples, even with a dispersion not completely ‘condensed’ between



**Fig. 2.** Representative  $^1\text{H}$  NMR spectra of coffee blends with resonance signals assigned: A shows the range of 0.0 to 10.0 ppm and B shows the expansion of the range of 1.8 to 4.4 ppm.

the groups. This shows that the specific data transformations in the multi-block methods are indeed interesting to highlight the contributions of the different blocks in the visualization of the data, which is generally not done by the PCA.

Despite the interesting differentiation, a detailed observation of the scores (Fig. 4B) shows that some samples of coffee blend 3 (rioysh/rio medium-dark roast), and mainly 1 (hard/rioysh medium-dark roast) and 2 (hard/rioysh dark roast), presented positive and negative scores. The samples of blend 1 and 2 belong to the same blend, but with different degrees of roasting. At the same time, samples from both blends (1, 2, and 3) are manufactured with rioysh coffee in common. As this study is carried out with real samples, these results provide two interesting observations. The first, in the case of coffee blends 1 and 2, which are similar but roasted under different conditions, the results show that if the roasting profile is not strictly followed by the professional roasters, the final blend profile can be compromised, leading to ambiguity in the final characteristics of the coffee blend, which is undesirable and very worrying in an industrial situation. The second point is about the presence of rioysh coffee in both blends. Since ComDim analysis is an

exploratory technique and is therefore unsupervised, it is possible to infer that the clusters formed reflects the true data structure and is not the result of inadequate data manipulation. In other words, this indicates that, despite the characteristics resulting from the roasting, the coffee blends can still bring chemical information from the initial cup profiles of coffees before blending, directly influencing the variation of the CC1 scores.

Regarding the loadings in Fig. 4C, which are used to display the variables that highlight the relationships among the data blocks, the coffee samples from blends 1 and 4 were related with  $^1\text{H}$  NMR resonance signals located on the negative side, NIR absorbance in the range 1400–1676 nm, color parameters ( $L^*$ ,  $a^*$ , and  $b^*$ ), and with all sensory properties (powder fragrance, drink aroma, acidity, bitterness, flavor, astringency, residual flavor, and overall quality), except with body sensory property. In contrast, coffee samples from blends 2 and 3 exhibited an opposite trend. These samples were related to  $^1\text{H}$  NMR resonance signals on the positive side, NIR absorbance in the 906 to 1400 nm range, moisture content, and body sensory property. The identification of these variables within each data block and their importance to

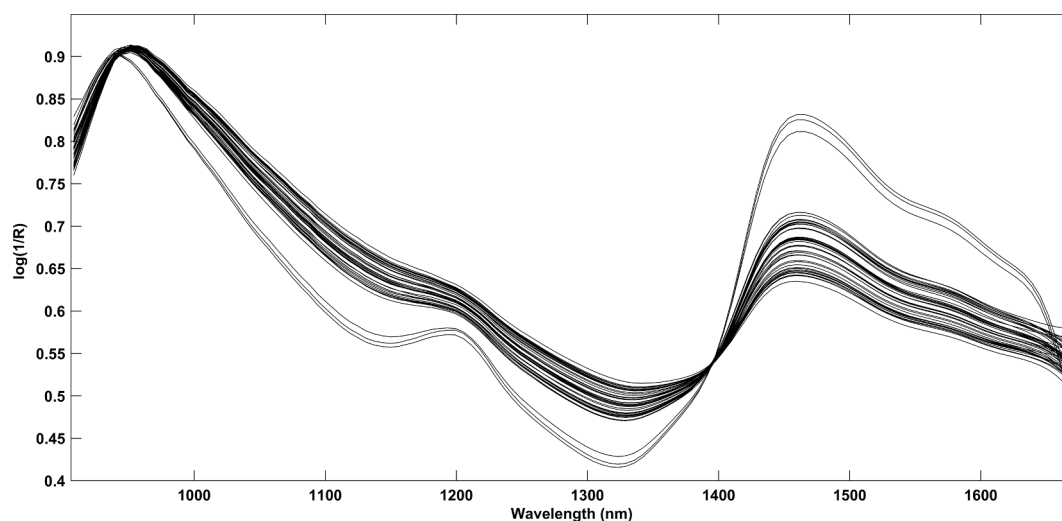


Fig. 3. NIR spectra of all roasted and ground coffee blends.

differentiate coffee blends are discussed in detail below.

Regarding the contribution of the metabolites, pointed out by the chemical shifts in ComDim analysis ( $^1\text{H}$  NMR loadings in Fig. 4C), it was feasible to highlight possible markers to characterize coffee blends. The  $^1\text{H}$  NMR loadings in Fig. 4C were expanded and shown in Fig. 5 with the assignment of resonance signal. From this plot (Fig. 5), the highlighted metabolites responsible for the differentiation of coffee blend samples were characterized by trigonelline ( $\delta$  4.45, 8.12, 8.84, and 9.13), formate ( $\delta$  8.49), caffeine ( $\delta$  3.29, 3.49, 3.89, and 7.85), chlorogenic acids ( $\delta$  5.37, 5.45, 6.31, 6.41, 6.88, 7.05, 7.13, and 7.61), choline ( $\delta$  3.21), lactate ( $\delta$  1.33), citrate and/or malate ( $\delta$  2.69), lipids ( $\delta$  0.84, 0.93, and 1.00), *N*-methylpyridinium ( $\delta$  4.41, 8.04, 8.57, and 8.81), quinic acids ( $\delta$  1.89, 2.01, 2.08, and 4.17) and *myo*-inositol ( $\delta$  3.33, 3.65, and 4.09). Among these metabolites, trigonelline, formate, caffeine, chlorogenic acids, choline, lactate, citrate/malate, and lipids were found to be markers for coffee blends with high cup quality and low degrees of roasting (coffee blends 1 and 4), while *N*-methylpyridinium, quinic acids, *myo*-inositol, and lipids were found as markers of blends with opposite characteristics: lower cup quality and high degrees of roasting (coffee blends 2 and 3).

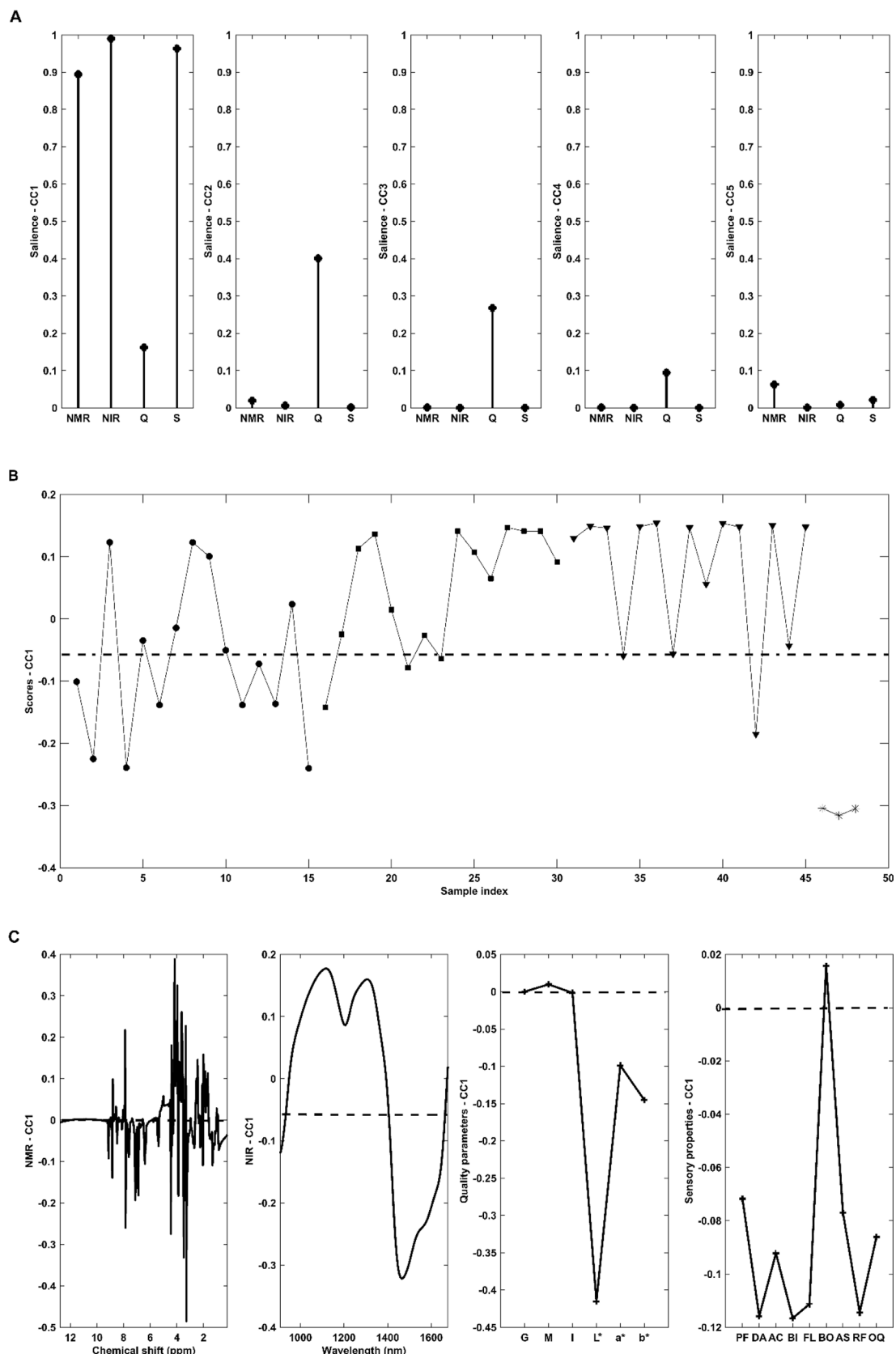
Regarding the loadings for portable NIR spectroscopy (Fig. 4B), indicated by the ComDim analysis, coffee blend samples with high cup quality and lower roasting degrees (coffee blends 1 and 4) were differentiated by strong absorptions in the region from 1400 to 1650 nm. The signals in this region can be attributed to carbohydrates, CGAs, and phenols vibrations, as well as signals of O—H first overtone from R—OH, and N—H first overtone from R—NH that can be related to lipids, proteins, caffeine, and water present in coffee (Barbin et al., 2014; Correia et al., 2018). On the other hand, coffee blend samples with lower cup quality and a high roasting degree (coffee blends 2 and 3) were characterized by strong absorptions in the region between 906 and 1400 nm, with two maximum vibrations, ranging from 1150 to 1200 nm and from 1200 to 1400 nm. The region of 900–1200 nm can be attributed to phenolic compounds and lipids in coffee (Correia et al., 2018; Monteiro et al., 2018; Ribeiro et al., 2011). This region also corresponds to the C—H third overtone, O—H first overtone from R—OH, N—H first overtone from R—NH, C—H third overtone from HC=CH, C—H second overtone from C=CH, C—H third and second overtones from CH<sub>3</sub>, and C—H third and second overtones from CH<sub>2</sub> (Barbin et al., 2014; Correia et al., 2018). The region between 1200 and 1400 nm can be attributed to caffeine, proteins, lipids, and carbohydrate vibrations, as well as other chemical assignments commons in coffee composition, such as the second overtone of C—H, the second overtone of C—H from CH<sub>2</sub>, the first overtone of C—H combination bands from CH<sub>3</sub>, and first overtone of

O—H from ArOH, CH<sub>2</sub>, and CH<sub>3</sub> (Barbin et al., 2014; Ribeiro et al., 2011).

Regarding the loadings for professional cupping (Fig. 4C), the ComDim analysis pointed out the powder fragrance, drink aroma, acidity, bitterness, flavor, astringency, residual flavor, and overall quality as the main sensory properties responsible for differentiating the coffee blends 1 and 4, while body sensory property was exclusively related to the coffee blends 2 and 3. When examining the average and range of sensory properties in Table S4 combined with the sensorial profile of coffee blends in Fig. S1 (both annexed in the Supplementary Material), most of the analyzed properties received notes very close to each other and overlap, except for overall quality (a synthesis of all properties) and body. Therefore, this indicates that, besides the overall quality, the beverage body is an important sensory marker to differentiate the final quality of coffee blends.

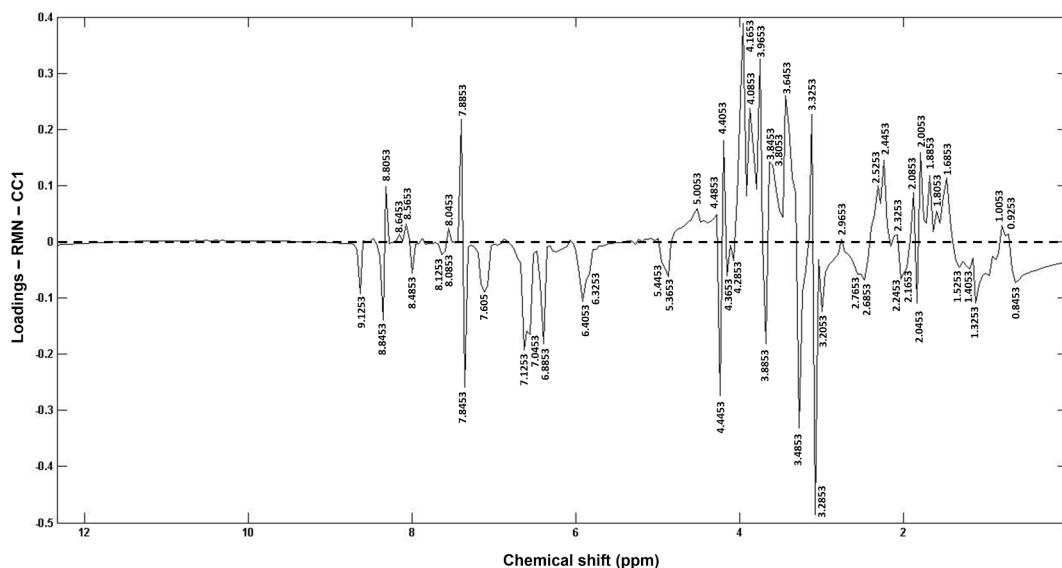
In an attempt to establish a relationship between the sensory properties and metabolite composition of coffee blends, which was supported by molecular information from portable NIR, the ComDim analysis demonstrated that quinic acids, *myo*-inositol, and mainly *N*-methylpyridinium and lipids were related to the property of the body, while trigonelline, formate, caffeine, chlorogenic acids, choline, lactate, citrate, and lipids were related to powder fragrance, drink aroma, acidity, bitterness, flavor, astringency, residual flavor, and overall quality. The relationship of metabolites with sensory properties is possible because greater and similar saliences proportions were found for the metabolic composition and sensory property data blocks, as well as for the NIR spectra (see Fig. 4A in CC1). As mentioned earlier, ComDim has the ability to reveal the relationships between data blocks when the saliences or “weights” of the data blocks are close. Previous studies (Craig et al., 2018; Farah et al., 2006; Franca et al., 2005; Ribeiro et al., 2011) have shown that among the sensory properties of the coffee beverage, the body is related to metabolites such as lipids and proteins. On the other hand, the overall quality is a combination of all sensory properties determined in cup quality evaluation (Ribeiro et al., 2011) and, therefore, it is expected that many metabolites are important for various attributes (powder fragrance, drink aroma, acidity, bitterness, flavor, astringency, and residual flavor) as well as for overall quality than when compared with just body sensory property. Thus, these relationships found between sensory properties and metabolic composition are consistent with these previous studies (Craig et al., 2018; Farah et al., 2006; Franca et al., 2005; Ribeiro et al., 2011).

The results obtained confirm the potential of the ComDim analysis to study multiple data blocks in the same samples and extract interesting information from the coffee blend profiles. In view of these results,



**Fig. 4.** ComDim results. (A) – saliences of the CC 1 to 5: where NIR = near-infrared spectroscopy; NMR = <sup>1</sup>H nuclear magnetic resonance, Q = quality parameters; S = sensory properties; (B) Scores of the CC 1: (●) Coffee blend 1; (■) Coffee blend 2; (▼) Coffee blend 3; (\*) Coffee blend 4; (C) Loadings of the CC 1, where G = granulometry; M = moisture content; I = infusion time; L = luminosity (quality parameters); PF = powder fragrance, DA = drink aroma, AC = acidity, BI = bitterness, FL = flavor; BO = body; AS = astringency; RF = residual flavor; OQ = overall quality (sensory properties).





**Fig. 5.**  $^1\text{H}$  NMR loadings expanded in ComDim analysis (block 1), i.e. zoom on  $^1\text{H}$  NMR loadings from Fig. 4C. Resonance signals of the main coffee metabolites highlighted by ComDim and located in negative side: trigonelline ( $\delta$  4.45, 8.12, 8.84, and 9.13), formate ( $\delta$  8.49), caffeine ( $\delta$  3.29, 3.49, 3.89, and 7.85), chlorogenic acids ( $\delta$  5.37, 5.45, 6.31, 6.41, 6.88, 7.05, 7.13, and 7.61), choline ( $\delta$  3.21), lactate ( $\delta$  1.33), citrate ( $\delta$  2.69), and lipids ( $\delta$  0.84); in positive side: *N*-methylpyridinium ( $\delta$  4.41, 8.04, 8.57, and 8.81); quinic acids ( $\delta$  1.89, 2.01, 2.08, and 4.17), *myo*-inositol ( $\delta$  3.33, 3.65, and 4.09), and lipids ( $\delta$  0.93 and 1.00).

spectral techniques, especially the portable NIR, offer a good alternative for monitoring product quality in the modern coffee industry. The main implication of this study for the industry is that the results obtained provide a way for evaluating and understanding the importance of the various methods and techniques typically used to evaluate coffee quality during its industrial production, demonstrating the potential of ComDim analysis to evaluate multiple data blocks originated from coffee samples.

#### 4. Conclusions

This work was the first study that evaluated coffee blends and their characteristics with a multi-block data analysis using a ComDim-based method in a comprehensive and realistic industrial perspective. The techniques and methods used in the industry represented the multiple blocks of investigated data. A data-fusion based PCA was also performed but was not effective in illustrating feasible results as ComDim. With a ComDim exploratory analysis applied to all data blocks, it was possible to demonstrate relationships between the information obtained from portable NIR,  $^1\text{H}$  NMR fingerprints, sensory properties, and the coffee roast degree, highlighting the importance of each block or criterion involved in the industrial quality control of coffee. The results showed that the coffee blends were differentiated based on their qualities and metabolite composition. Coffee blends with high cup quality and lower roasting degree were generally differentiated from those with opposite characteristics: lower cup quality and high roasting degree. Therefore, the results achieved confirm the ability, applicability, reliability and feasibility of multi-block data analysis in the field of coffee science.

#### CRediT authorship contribution statement

**Michel Rocha Baqueta:** Data curation, Formal analysis, Investigation, Methodology, Writing - original draft. **Aline Coqueiro:** Conceptualization, Supervision, Formal analysis, Investigation, Data curation, Writing - review & editing. **Paulo Henrique Março:** Conceptualization, Investigation, Methodology, Formal analysis. **Ferruccio Poli:** Investigation, Methodology. **Patrícia Valderrama:** Conceptualization, Supervision, Funding acquisition, Formal analysis, Investigation, Data curation, Writing - review & editing.

#### Declaration of Competing Interest

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

#### Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. Patrícia Valderrama thanks Fundação Araucária (process 033/2019). We thank also the CAMulti - UTFPR Campo Mourão for providing equipment and infrastructure.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.foodchem.2021.129618>.

#### References

- ABIC. (2018). Norma de Qualidade Recomendável e Boas Práticas de Fabricação de Cafés Torrados em Grãos e Cafés Torrados e Moídos. Associação Brasileira Da Indústria de Café, 1–38. Retrieved from <http://abic.com.br/src/uploads/2017/07/2.8.1-Norma-de-qualidade-PQC.pdf>.
- Assis, C., Pereira, H. V., Amador, V. S., Augusti, R., de Oliveira, L. S., & Sena, M. M. (2019). Combining mid-infrared spectroscopy and paper spray mass spectrometry in a data fusion model to predict the composition of coffee blends. *Food Chemistry*, 281 (December 2018), 71–77. [10.1016/j.foodchem.2018.12.044](https://doi.org/10.1016/j.foodchem.2018.12.044).
- Baqueta, M. R., Coqueiro, A., Março, P. H., & Valderrama, P. (2020). Quality Control Parameters in the Roasted Coffee Industry: A Proposal by Using MicroNIR Spectroscopy and Multivariate Calibration. *Food Analytical Methods*, 13(1), 50–60. <https://doi.org/10.1007/s12161-019-01503-w>.
- Baqueta, M. R., Coqueiro, A., & Valderrama, P. (2019). Brazilian Coffee Blends: A Simple and Fast Method by Near-Infrared Spectroscopy for the Determination of the Sensory Attributes Elicited in Professional Coffee Cupping. *Journal of Food Science*, 84(6), 1247–1255. <https://doi.org/10.1111/1750-3841.14617>.
- Baqueta, M. R., Do Prado Silva, J. T., Moya Moreira, T. F., Canesin, E. A., Gonçalves, O. H., Dos Santos, A. R., ... Leimann, F. V. (2017). Extração e caracterização de compostos do resíduo vegetal casca de café. *Brazilian Journal of Food Research*, 8(2), 68. <https://doi.org/10.3895/rebrapa.v8n2.6887>.
- Barbin, D. F., Felicio, A. L. d. S. M., Sun, D.-W., Nixdorf, S. L., & Hirooka, E. Y. (2014). Application of infrared spectral techniques on quality and compositional attributes of coffee: An overview. *Food Research International*, 61, 23–32. <https://doi.org/10.1016/j.foodres.2014.01.005>.

- Bertone, E., Venturello, A., Giraudo, A., Pellegrino, G., & Geobaldo, F. (2016). Simultaneous determination by NIR spectroscopy of the roasting degree and Arabica/Robusta ratio in roasted and ground coffee. *Food Control*, 59, 683–689. <https://doi.org/10.1016/j.foodcont.2015.06.055>.
- Bouveresse, D.-J.-R., Pinto, R. C., Schmidtke, L. M., Locquet, N., & Rutledge, D. N. (2011). Identification of significant factors by an extension of ANOVA-PCA based on multi-block analysis. *Chemometrics and Intelligent Laboratory Systems*, 106(2), 173–182. <https://doi.org/10.1016/j.chemolab.2010.05.005>.
- Brazil. (2003). Instrução Normativa no 8, de 11 de junho de 2003.
- Cariou, V., Jouan-Rimbaud Bouveresse, D., Qannari, E. M., & Rutledge, D. N. (2019). ComDim Methods for the Analysis of Multiblock Data in a Data Fusion Perspective. In *Data Handling in Science and Technology* (Vol. 31), 10.1016/B978-0-444-63984-4.00007-7.
- Combes, M. C., Joët, T., & Lashermes, P. (2018). Development of a rapid and efficient DNA-based method to detect and quantify adulterations in coffee (Arabica versus Robusta). *Food Control*, 88, 198–206. <https://doi.org/10.1016/j.foodcont.2018.01.014>.
- Consonni, R., Polla, D., & Cagliani, L. R. (2018). Organic and conventional coffee differentiation by NMR spectroscopy. *Food Control*, 94, 284–288. <https://doi.org/10.1016/J.FOODCONT.2018.07.013>.
- Cordella, C. B. Y., & Bertrand, D. (2014). SAISIR: A new general chemometric toolbox. *TrAC - Trends in Analytical Chemistry*, 54, 75–82. <https://doi.org/10.1016/j.trac.2013.10.009>.
- Correia, R. M., Tosato, F., Domingos, E., Rodrigues, R. R. T., Aquino, L. F. M., Filgueiras, P. R., ... Romão, W. (2018). Portable near infrared spectroscopy applied to quality control of Brazilian coffee. *Talanta*, 176, 59–68. <https://doi.org/10.1016/j.talanta.2017.08.009>.
- Craig, A. P., Botelho, B. G., Oliveira, L. S., & Franca, A. S. (2018). Mid infrared spectroscopy and chemometrics as tools for the classification of roasted coffees by cup quality. *Food Chemistry*, 245, 1052–1061. <https://doi.org/10.1016/J.FOODCHEM.2017.11.066>.
- Dias, R. C. E., Valderrama, P., Março, P. H., dos Santos Scholz, M. B., Edelmann, M., & Yeretizian, C. (2018). Quantitative assessment of specific defects in roasted ground coffee via infrared-photoacoustic spectroscopy. *Food Chemistry*, 255, 132–138. <https://doi.org/10.1016/J.FOODCHEM.2018.02.076>.
- dos Santos, H. D., Alvarenga, Y. A., & Boffo, E. F. (2020). 1H NMR metabolic fingerprinting of Chapada Diamantina/Bahia (Brazil) coffees as a tool to assessing their qualities. *Microchemical Journal*, 152, 104293. <https://doi.org/10.1016/j.microc.2019.104293>.
- El Ghaziri, A., Cariou, V., Rutledge, D. N., & Qannari, E. M. (2016). Analysis of multiblock datasets using ComDim: Overview and extension to the analysis of (K + 1) datasets. *Journal of Chemometrics*, 30(8), 420–429. <https://doi.org/10.1002/cem.2810>.
- Farah, A., Monteiro, M. C., Calado, V., Franca, A. S., & Trugo, L. C. (2006). Correlation between cup quality and chemical attributes of Brazilian coffee. *Food Chemistry*, 98(2), 373–380. <https://doi.org/10.1016/J.FOODCHEM.2005.07.032>.
- Ferreira, M. M. C. (2015). Quimiometria: conceitos, métodos e aplicações (1st ed.). Retrieved from <https://books.google.com.br/books?hl=pt-BR&lr=&id=a5OnDwAAQBAJ&oi=fnd&pg=PA264&dq=+Quimiometria:+conceitos,+métodos+e+aplicações&ots=UKGDAGB2QD&sig=0XNPYn8py4HljY1b2e7wwh-37dQ>.
- Franca, A. S., Mendonça, J. C. F., & Oliveira, S. D. (2005). Composition of green and roasted coffees of different cup qualities. *LWT - Food Science and Technology*, 38(7), 709–715. <https://doi.org/10.1016/J.LWT.2004.08.014>.
- Geladi, P., MacDougall, D., & Martens, H. (1985). Linearization and Scatter-Correction for Near-Infrared Reflectance Spectra of Meat. *Applied Spectroscopy*, 39(3), 491–500. <https://doi.org/10.1366/0003702854248656>.
- Giacalone, D., Fosgaard, T. R., Steen, I., & Münchow, M. (2016). “Quality does not sell itself”: Divergence between “objective” product quality and preference for coffee in naïve consumers. *British Food Journal*, 118(10), 2462–2474. <https://doi.org/10.1108/BFJ-03-2016-0127>.
- Jumhawan, U., Putri, S. P., Yusianto, B., & T., & Fukusaki, E. (2016). Quantification of coffee blends for authentication of Asian palm civet coffee (Kopi Luwak) via metabolomics: A proof of concept. *Journal of Bioscience and Bioengineering*, 122(1), 79–84. <https://doi.org/10.1016/j.jbiosc.2015.12.008>.
- Loudiyi, M., Rutledge, D. N., & Ait-Kaddour, A. (2018). ComDim for explorative multi-block data analysis of Cantal-type cheeses: Effects of salts, gentle heating and ripening. *Food Chemistry*, 264, 401–410. <https://doi.org/10.1016/J.FOODCHEM.2018.05.039>.
- Martin, Jean-Charles, Maillot, Matthieu, Mazerolles, Gérard, Verdu, Alexandre, Lyan, Bernard, Migné, Carole, ... Rutledge, Douglas N. (2015). Can we trust untargeted metabolomics? Results of the metabo-ring initiative, a large-scale, multi-instrument inter-laboratory study. *Metabolomics*, 11(4), 807–821. <https://doi.org/10.1007/s11306-014-0740-0>.
- Milani, M. I., Rossini, E. L., Catelani, T. A., Pezza, L., Toci, A. T., & Pezza, H. R. (2020). Authentication of roasted and ground coffee samples containing multiple adulterants using NMR and a chemometric approach. *Food Control*, 112(November 2019), 107104. <https://doi.org/10.1016/j.foodcont.2020.107104>.
- Mishra, P., Roger, J. M., Rutledge, D. N., Biancolillo, A., Marini, F., Nordon, A., & Jouan-Rimbaud-Bouveresse, D. (2020). MBA-GUI: A chemometric graphical user interface for multi-block data visualisation, regression, classification, variable selection and automated pre-processing. *Chemometrics and Intelligent Laboratory Systems*, 205(7), Article 104139. <https://doi.org/10.1016/j.chemolab.2020.104139>.
- Monteiro, Pablo Inocêncio, Santos, Jânio Sousa, Alvarenga Brizola, Vitor Rafael, Pasini Deolindo, Carolina Turnes, Koot, Alex, Boerrigter-Eenling, Rita, ... Granato, Daniel (2018). Comparison between proton transfer reaction mass spectrometry and near infrared spectroscopy for the authentication of Brazilian coffee: A preliminary chemometric study. *Food Control*, 91, 276–283. <https://doi.org/10.1016/j.foodcont.2018.04.009>.
- Qannari, E. M., Wakeling, I., Courcoux, P., & MacFie, H. J. (2000). Defining the underlying sensory dimensions. *Food Quality and Preference*, 11(1–2), 151–154. [https://doi.org/10.1016/S0950-3293\(99\)00069-5](https://doi.org/10.1016/S0950-3293(99)00069-5).
- Ribeiro, J. S., Ferreira, M. M. C., & Salva, T. J. G. (2011). Chemometric models for the quantitative descriptive sensory analysis of Arabica coffee beverages using near infrared spectroscopy. *Talanta*, 83(5), 1352–1358. <https://doi.org/10.1016/J.TALANTA.2010.11.001>.
- Ríos-Reina, R., Callejón, R. M., Savorani, F., Amigo, J. M., & Cocchi, M. (2019). Data fusion approaches in spectroscopic characterization and classification of PDO wine vinegars. *Talanta*, 198, 560–572. <https://doi.org/10.1016/J.TALANTA.2019.01.100>.
- Rosa, Larissa Naida, de Figueiredo, Luana Caroline, Bonafé, Elton Guntendorfer, Coqueiro, Aline, Visentainer, Jesuí Vergílio, Março, Paulo Henrique, ... Valderrama, Patrícia (2017). Multi-block data analysis using ComDim for the evaluation of complex samples: Characterization of edible oils. *Analytica Chimica Acta*, 961, 42–48. <https://doi.org/10.1016/j.aca.2017.01.019>.
- Savitzky, A., & Golay, M. J. E. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, 36(8), 1627–1639. <https://doi.org/10.1021/ac60214a047>.
- Souard, F., Delporte, C., Stoffelen, P., Thévenot, E. A., Noret, N., Dauvergne, B., ... Stévigny, C. (2018). Metabolomics fingerprint of coffee species determined by untargeted-profiling study using LC-HRMS. *Food Chemistry*, 245(September 2017), 603–612. <https://doi.org/10.1016/j.foodchem.2017.10.022>.
- Toci, A. T., de Moura Ribeiro, M. V., de Toledo, P. R. A. B., Boralle, N., Pezza, H. R., & Pezza, L. (2018). Fingerprint and authenticity roasted coffees by 1H-NMR: The Brazilian coffee case. *Food Science and Biotechnology*, 27(1), 19–26. <https://doi.org/10.1007/s10068-017-0243-7>.
- Wei, F., Furihata, K., Miyakawa, T., & Tanokura, M. (2014). A pilot study of NMR-based sensory prediction of roasted coffee bean extracts. *Food Chemistry*, 152, 363–369. <https://doi.org/10.1016/J.FOODCHEM.2013.11.161>.
- Wermelinger, T., D'Ambrosio, L., Klopprogge, B., & Yeretizian, C. (2011). Quantification of the robusta fraction in a coffee blend via raman spectroscopy: Proof of principle. *Journal of Agricultural and Food Chemistry*, 59(17), 9074–9079. <https://doi.org/10.1021/jf201918a>.