POLITECNICO DI TORINO
Repository ISTITUZIONALE

Monitoring Chemical Changes of Coffee Beans During Roasting Using Real-time NIR Spectroscopy and Chemometrics

Original
Monitoring Chemical Changes of Coffee Beans During Roasting Using Real-time NIR Spectroscopy and Chemometrics / Grassi, S., Giraudo, A., Novara, C., Cavallini, N., Geobaldo, F., Casiraghi, E., Savorani, F.. - In: FOOD ANALYTICAL METHODS. - ISSN 1936-9751. - ELETTRONICO. - (2023). [10.1007/s12161-023-02473-w]

Availability:
This version is available at: 11583/2977230 since: 2023-03-20T09:42:57Z

Publisher:
Springer

Published
DOI:10.1007/s12161-023-02473-w

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

18 November 2026

Food Analytical Methods
https://doi.org/10.1007/s12161-023-02473-w

Monitoring Chemical Changes of Coffee Beans During Roasting Using Real-time NIR Spectroscopy and Chemometrics

Silvia Grassi[1] · Alessandro Giraudo[2] · Chiara Novara[2] · Nicola Cavallini[2] · Francesco Geobaldo[2] · Ernestina Casiraghi[1] · Francesco Savorani[2]

Received: 21 November 2022 / Accepted: 6 March 2023
© The Author(s) 2023

Abstract
Variations occurring in coffee beans during roasting are ascribable to several chemical-physical phenomena: to quickly track the whole process and to ensure its reproducibility, a process analytical technology (PAT) approach is needed.
In this study, a method combining in-line Fourier transform near-infrared (FT-NIR) spectroscopy and chemometric modelling was investigated to get real-time and practical knowledge about the roasting effects on coffee's chemical-physical composition. In-line spectra were acquired by inserting a NIR probe into a laboratory coffee roaster, running twenty-four roasting experiments, planned spanning different coffee species (Arabica and Robusta), four roasting temperature settings (TS1–TS4) and times (650–1580 s).
Multivariate curve resolution-alternate least squares (MCR-ALS) was used to model the chemical-physical changes occurring during the roasting process, and information about maximum rate, acceleration and deceleration of the process was obtained, also highlighting potential effects due to the different roasting temperatures and coffee varieties.
The proposed approach provides the groundwork for direct real-time implementation of rapid, non-invasive automated monitoring of the roasting process at industrial scale.

Keywords Coffee · Roasting process · In-line analysis · NIR spectroscopy · MCR-ALS

Introduction

Coffee is one of the most popular beverages since millions of people all over the world consume it several times a day. Roasting is an essential step in the coffee production process, as it enables the development of colour, aroma, and flavour, which are crucial for coffee quality. During roasting, the coffee beans undergo a steady weight loss, due to initial dehydration and then to a complex set of chemical and physical processes (Iaccheri et al. 2015). At the same time, the beans undergo a significant increase in volume, essentially caused by internal formation of gas (CO_2 and other thermal decomposition products) (Franca et al. 2005). During roasting, a large number of chemical transformations take place simultaneously, leading to the formation of thousands of different molecules (De Maria et al. 1996). The chemical composition of roasted beans is related to the variety and geographical origin of green coffee, and it has a relevant influence on the sensorial properties of the beverage. Furthermore, the roasting conditions and technological, agronomic and physical parameters can deeply affect the final coffee quality (Esteban-Díez et al. 2007). Thus, the roasting process must be controlled accurately to meet the consumers' expectations, which are accustomed to a constant and consistent level of coffee quality.

For this reason, the assessment of the optimal roasting degree during the roasting process represents a powerful screening tool to quickly evaluate the overall performance of the process, and to ensure its reproducibility. This important indicator is closely related to the sensorial properties of the final coffee beverage, and it was found to be highly

Silvia Grassi and Alessandro Giraudo contributed equally to this work.

✉ Nicola Cavallini
nicola.cavallini@polito.it

[1] Department of Food, Environmental and Nutritional Sciences, University of Milan, Via G. Celoria 2, 20133 Milan, Italy

[2] Department of Applied Science and Technology, Polytechnic of Turin, Corso Duca Degli Abruzzi 24, 10129 Turin, Italy

Published online: 18 March 2023

affected not only by both the characteristics and the species of the green coffee beans, but also by the roasting process conditions (Andueza et al. 2007). The coffee roasting degree has been experimentally assessed using different analytical approaches, mainly based on off-line analysis of coffee batches, i.e. gathering this information only at the end of the process. In this context, earlier research studies focused on detecting the variations in the chemical composition profile of the gas developed during roasting using gas chromatography-flame ionization detector (GC-FID) systems (Dutra et al. 2001), whereas in more recent studies, the volatile component profile of the coffee beans was analysed under different process temperatures using solid-phase microextraction (SPME) and gas chromatography-mass spectrometry (GC–MS) (Franca et al. 2009), or by monitoring the chemical composition of the roasting off gases by online mass spectrometry (Heide et al. 2020; Czech et al. 2020). All these techniques remain relatively distant from real-time and on/in-line process control setups.

During the last decades, the use of spectroscopic approaches, mainly based on near-infrared (NIR) spectroscopy, has proven to be undoubtedly advantageous with respect to other classical analytical techniques, due to the reduced time between performing the analyses and obtaining the results, as well as to the reduced or even unnecessary sample pre-treatment. Nowadays, NIR spectroscopy associated with chemometrics is a widespread and well-known methodology already adopted to monitor crucial quality parameters of food products and processes, leading to the implementation of process analytical technology (PAT) approaches in the food industries (Grassi and Alamprese 2018). In the framework of coffee roasting, NIR spectroscopy has been successfully applied to predict important coffee quality parameters and to determine the roasting degree itself. Different off-line applications have been proposed to predict sensory attributes of espresso coffees from the NIR spectra of roasted coffee beans (Esteban-Díez et al. 2004) and to estimate roasting colour and caffeine content (Pizarro et al. 2007). As far as it concerns roasting, different approaches have been investigated for predicting relevant process variables such as weight loss, density and moisture (Alessandrini et al. 2008; Shan et al. 2015), roasting degree (Bertone et al. 2016; Pires et al. 2021) and the effects of different roasting conditions on coffee samples having different geographical origins (De Luca et al. 2016). NIR has been also investigated, using on-line or even in-line acquisition probes, for the quantification of acidity (Santos et al. 2016a), sugar content and colour (Santos et al. 2016b) and antioxidant properties (Catelani et al. 2017) of coffee beans during the roasting process. An interesting study by Catelani et al. (Catelani et al. 2018) evaluated the effectiveness of an approach based on the combination of NIR spectroscopy and multivariate data analysis for the early detection of roasting process disturbances. Moreover, in situ NIR spectroscopy was successfully applied by Yergenson and Aston for the real-time determination of coffee roasting degree based on the prediction of first and second crack events during the roasting process (Yergenson and Aston 2020b, a), by combining the chemical information of the NIR spectra with the audio recordings of the cracking events, thus not continuously tracking the roasting process.

All the above-mentioned research studies clearly highlighted that the changes occurring during roasting are ascribable to many different chemical and physical parameters, which are potentially correlated with each other and can be determined using fast and non-invasive analytical techniques like NIR spectroscopy.

In the present study, the potential of an approach based on in-line Fourier transform near-infrared (FT-NIR) spectroscopy coupled with chemometric modelling was investigated at laboratory scale to track the coffee roasting evolution and to assess the process performance in real-time from an overall viewpoint, i.e. by considering many different chemical-physical parameters at the same time.

The in-line FT-NIR spectra were acquired on the entire coffee beans by inserting a NIR probe into a laboratory coffee roaster. Several roasting experiments were performed considering different coffee species, roasting temperature settings and roasting time durations. The FT-NIR spectra were modelled using principal component analysis (PCA) and multivariate curve resolution-alternating least squares (MCR-ALS) (de Juan and Tauler 2006), with the aim of obtaining structured information about the roasting progress. Furthermore, three kinetics critical parameters of the process (maximum rate, acceleration and deceleration) were derived from the MCR-ALS concentration profiles, and the potential effects of roasting temperature or coffee variety on these kinetic parameters were investigated and discussed.

The final aim of this study was to explore the reliability of a roasting monitoring tool based on NIR technology to be implemented in the coffee roasting chain for automating the real-time control of the roaster's parameters.

Materials and Methods

Setup of the Roasting System

The roasting experiments were conducted using a laboratory coffee roaster (Sample PROBAT PRZ 2/4/6 Barrel Roaster, CoffeeTec, Germany) kindly provided by a leading Italian coffee industry. The roaster was mounted on a wood base, and it was electrically powered. It consisted of a cylindrical rotating drum with a circular window on the front end and a hopper to load the coffee beans on the back end. The drum had a dedicated air temperature display to allow simple adjustments to the heat dial. Temperature control was

realized by means of a PID controller, which regulated two resistors positioned on the bottom of the drum. During the roasting process, the cylinder was kept in rotation to ensure even roasting for all the coffee beans.

Preliminary roasting tests were performed to identify the best setup for the in-line acquisition of the NIR spectra during the process. In this context, some relevant parameters had to be considered. First, the appropriate stream of air to flux inside the roaster during the process had to be determined. This parameter could be manually regulated in each drum by means of a lever. Its adjustment is important to avoid the accumulation of vapour and combustion fumes, which normally develop during the roasting process and that can deeply affect the quality of the acquired NIR spectra. As a result of the preliminary tests, the air flow was set at a constant level of medium intensity to remove the gaseous components while avoiding, at the same time, excessive cooling of the coffee beans during the process.

A second parameter that had to be accurately controlled is the air temperature: four different temperature settings were considered (hereafter referred to as TS1, TS2, TS3 and TS4). An optical temperature probe (TS Temperature Sensors, Micronor LLC, USA) was employed to measure the actual temperature of the coffee beans in real time (one acquisition per second) throughout the roasting process.

In-line NIR Acquisition System

A total of 1636 in-line NIR spectra were collected during 24 roasting experiments. The NIR spectra were acquired in diffuse reflectance mode using an "Emission Head" NIR probe coupled to a MATRIX-F FT-NIR spectrometer (Bruker Optics, Germany). The Emission Head allowed collecting the NIR spectra by placing the probe on a stand near the open end of the roaster. An angle of incidence of 45° downwards between the probe and the window of the roaster was found to be the optimal solution to properly focus the NIR light beam on the coffee beans during the roasting process. The NIR spectra were acquired with a resolution of 8 cm^{-1}, 64 co-added scans and considering the wavenumber range between 10,265 and 4135 cm^{-1} (974–2418 nm). The FT-NIR spectrometer was controlled via OPUS software (version 7.5, Bruker Optics, Germany). The considered NIR acquisition system allowed collecting one spectrum every 15 s. Before starting the NIR spectra acquisition during each roasting experiment, a background spectrum was acquired using a 100% reflectance reference material (PTFE Teflon).

Coffee Samples and Roasting Conditions

The coffee samples treated in this study originated from the same harvesting period and were kindly provided by a leading Italian coffee industry. The coffee beans showed worldwide geographical origins, and both the two most relevant commercial species, i.e. Arabica and Robusta, were considered. The roasting experiments were carried out within 48 h from the sample delivery and were planned considering an experimental procedure covering coffee species and roasting temperature levels.

Concerning the coffee species, all experiments were performed considering Arabica and Robusta separately. Four levels of increasing roasting temperature (from TS1–TS4) were considered to span a large enough range of temperature values with respect to the common industrial setup (200–220 °C). Finally, the roasting process lasted from a minimum of 650 to a maximum of 1580 s, depending on the temperature setting. Three different roasting batches were performed considering the Arabica and Robusta species separately at each temperature level, thus for a total of 24 experiments (2 species × 4 temperatures × 3 analytical replicates).

The actual average temperature profiles in the roaster along the process recorded by the optical temperature probe are reported in Figure S1 (strong temperature drops occurred during the roaster loading and discharge steps, marking the beginning and the end of the roasting process, and were excluded from the charts for clarity). All the temperature profiles were featured by an initial rapid temperature increase after coffee loading, followed by the approaching of a plateau, that was however reached only in the case of the long-lasting TS1 tests. The temperature always went over 210 °C during the tests, even if after longer times for the low roasting temperature experiments. Moreover, the maximum temperature was significantly higher for the TS3 and TS4 (~ 231 °C and 238 °C at the end time of the shortest roasting processes of the set, respectively) compared to the TS1 and TS2 (~ 218 °C and 212 °C, respectively), despite in the former cases the process was usually stopped earlier. Finally, good reproducibility of the measured temperature profiles was achieved, as shown by the low standard deviation calculated for the average profiles (Figure S1), pointing out that the coffee batches roasted at the same temperature level were subjected to very similar operative conditions.

Execution of the Roasting Experiments

Before beginning each roasting experiment, 120 g of green (i.e. unroasted) coffee beans was weighed (VWR scale 611–2602, accuracy of ± 0.01 g; VWR, Radnor), while their moisture analysis was carried out by means of a thermobalance (RADWAG MAC 210/NH; Radom, Poland). The same analysis was performed after roasting to evaluate the sample weight loss.

When the temperature set on the roaster was reached, the coffee sample was introduced into one of the roaster's drums. At the same time, the temperature and NIR probes, previously directed towards the inlet lid of the roaster drum,

started recording temperature data and spectra. When the process ended, according to the scheduled time of the corresponding experiment, the coffee was discharged into a container and cooled down by a stream of cool air, while the acquisition of the NIR spectra and the coffee temperature were promptly halted, and the data saved.

The coffee colour was determined on 25–30 g of roasted coffee sample using the single-beam Color Test II (Neuhaus Neotec, Germany), which performs a reflectance acquisition at 900 nm, with a ± 0.01 Arbitrary Unit (A.U.) accuracy. The same instrument was used to determine the colour of ground roasted coffee in our previous work (Bertone et al. 2016).

The details of the randomized roasting experiments, in terms of coffee species, roasting conditions, colour, initial moisture and weight loss, are reported in the Supplementary Materials in Table S1.

Multivariate Data Analysis

Chemometric elaboration of the NIR spectra was performed under MATLAB environment (ver. 9.2, The Mathworks, Inc., Natick, USA) using appropriate software packages, i.e. the "PLS_Toolbox" (ver. 8.5, Eigenvector Research, Inc., USA), the "MCR-ALS 2.0 Toolbox" (freely downloadable from https://mcrals.wordpress.com/download/mcr-als-2-0-toolbox last access 24/02/2023) and additional in-house generated MATLAB scripts. The 1636 in-line NIR spectra that were collected during the 24 roasting experiments were organized in a single dataset, where the spectra belonging to each roasting experiment were labelled and paired with all the additional information related to the process conditions and the coffee samples, as listed in Table S1.

Pre-processing of NIR Spectra and Exploratory Data Analysis

Different data pre-processing techniques were tested to remove potential artefacts from the NIR spectra, to correct nonlinear behaviours and to improve the signal-to-noise ratio. In this context, the "one-dimensional median filter" ("medfilt1" MATLAB function) algorithm was applied to remove, when needed, spikes from the acquired NIR spectra. Subsequently, the performances of the Savitzky-Golay smoothing algorithm were investigated, alone or in combination with the standard normal variate (SNV) or multiple scattering correction (MSC) algorithms. Regardless of the applied pre-processing, all the data were subjected to mean centering prior to any multivariate data analysis.

The exploratory data analysis was carried out by principal component analysis (PCA) (Wall et al. 2005). The optimal number of principal components (PCs) to be retained in the PCA model was chosen from the analysis of the corresponding scree plot (Bro and Smilde 2014). First, PCA was used to identify and remove the outliers, which mainly corresponded to initial spectra recorded before the coffee roaster was completely loaded. Subsequently, different PCA models were built to gather a preliminary overview of spectral variability in the multivariate space. In this last case, all the combinations of the above-mentioned pre-processing methods, coupled with mean centering, were investigated.

MCR-ALS Modelling

Multivariate curve resolution-alternating least squares (MCR-ALS) (de Juan and Tauler 2006) was employed to get a deeper insight into the evolution of the roasting process at each considered roasting temperature. MCR is a decomposition technique which allows extracting deconvoluted "pure" spectral profiles and their relative concentrations in each sample, using an ALS procedure. Once the number of components to be modelled is defined, the output of MCR is a set of resolved spectra (one for each component) and a set of relative concentrations (one for each component). The pure spectral profiles can be inspected to identify all relevant signals, thus allowing for proper chemical interpretation of each modelled component. The relative concentrations can be inspected to find similarities and differences among the samples, and if these correspond to specific time points (like in the present study) the evolution in time can be easily inspected.

In this context, the whole "clean" spectral dataset (i.e. after removing the outliers) was split into four subsets, each one including the experiments performed using both the Arabica and the Robusta coffee samples, and at the same roasting temperature level. On each subset, a separate MCR-ALS model was thus obtained.

The models were initially developed by considering the whole spectral range (10,265–4135 cm^{-1}). However, to further evaluate the performance of this modelling approach and in view of the industrial use of simplified NIR instruments with a reduced spectral range, models based on the selection of a limited number of spectral variables were also implemented. In detail, a reduced spectral range (range 1, 6120–4961 cm^{-1}) was considered according to the regions identified by Catelani et al. (Catelani et al. 2018) as the most suitable to monitor and control the roasting process. Moreover, a second spectral range was tested (range 2, 9777–7417 cm^{-1}), which is mainly related to the physical properties of the coffee beans, such as their size, shape and density, as suggested by Catelani et al. (Catelani et al. 2018).

In all cases, the optimal number of MCR components to be used in the model was determined by means of PCA. An initial estimate of the pure spectral matrix being calculated, the non-negativity constraint was imposed on both the concentration (**C**) and the profile (**S**) matrices using the fast non-negative least squares algorithm ("fnnls" option in

the toolbox (Bro and De Jong 1997)), whereas normalization of the **S** matrix was performed using the "Euclidean" criterion (de Juan and Tauler 2006). The ALS optimization was performed considering 100 iterations and 0.1 as the convergence criterion, while the model performance was evaluated in terms of fitting error (lack of fit, LOF) of explained variance and standard deviation of the residuals.

Kinetic Fitting of the MCR-ALS Concentration Profiles

The concentration profiles of each model were fitted as a function of roasting time, using the sigmoid Eq. (1) implemented in Table Curve software (v. 4.0, Jandel Scientific, San Rafael, CA, USA) as reported by Grassi et al. (Grassi et al. 2013):

$$y = a + \text{bexp}\left(-exp\left(-\left(\frac{(x - d1n(1n(2)) - c)}{d}\right)\right)\right) \quad (1)$$

The first and second derivatives of the calculated sigmoid functions were used to identify the kinetic critical points, i.e. time-related to the maximum rate and acceleration and deceleration of the roasting phenomena. Two-way univariate analysis of variance (ANOVA) was performed on the kinetic indexes obtained for all the models, to assess the influence of roasting temperatures and the differences among the spectral ranges used in the MCR-ALS models. In both cases, coffee bean variety was considered as the second independent factor, together with the other factor, i.e. temperature or spectral range. Furthermore, PCA models were built to better visualise the effect of the temperature and the variety on the kinetic parameters derived from the MCR-ALS models developed for each spectral range considered, as reported in Fig. 5.

Results and Discussion

Exploratory Data Analysis

The PCA model developed on the mean-centred NIR spectra led to successful detection of the outliers (mainly spectra acquired when the toaster was not completely loaded) as shown in Fig. 1a, with colour coding reflecting the roasting temperature (TS1–TS4). The coffee-related samples are grouped into a very compact cluster located on the top-left section of the plot, whereas the outliers appear to be randomly scattered within the fourth quadrant. Regarding the coffee-related cluster, the sample distribution along PC2 can be ascribed to the NIR spectra variability during the roasting process. To better explain the samples' distribution in the score plot, the loadings of PC1 and PC2 are reported in Fig. 1b. It was found that for both components, most of the loading values resulted different from zero, meaning that almost all the spectral regions contributed to defining the observed variability.

To confirm the outliers' identification, the Hotelling T^2 statistics was inspected. This parameter corresponds to Student's *t*-statistic for multivariate data and measures the distance of each sample within the model. This approach is commonly used in the frame of the multivariate statistical process control (MSPC) to get an earlier detection of extreme, atypical and not well-modelled sampled values leading to strange process behaviours (Li Vigni et al. 2013; Catelani et al. 2018). Figure 1c reports the Hotelling T^2 statistic calculated for each collected spectrum, grouped according to the experimental run and ordered according to the roasting time passing. The plot, together with the inset figure, clearly shows how the spectra collected immediately at the beginning and at the end of each roasting experiment have higher T^2 values than the remaining spectra belonging to the same experiment (with 95% of confidence limit). Similar results were also obtained from the roasting experiments performed with Robusta coffee samples (data not shown). Therefore, the Hotelling T^2 statistic demonstrated to be a satisfactory tool for the rapid and automatic detection of the spectra to be considered outliers. For a graphical example of problematic spectra at the beginning and the end of the roasting process, please see Figure S2.

After removing the outliers, only the 1501 coffee-related spectra were subsequently explored with PCA. The best model was obtained by applying the following pre-processing methods: median filter (order 30), Savitzky-Golay smoothing (window width 11, polynomial order 2, derivative order 0) and mean centering.

A graphical representation of the PCA results is given in Fig. 2. The score plot of the first two PCs coloured according to the roasting time is reported in Fig. 2a, and it can be deduced that PC1 (which corresponds to the largest amount of the dataset's variability 84.63%) is related to the roasting time. The spectra collected on coffee samples at the beginning of the roasting process are located at positive values on PC1, and by moving towards more and more negative PC1 score values can be interpreted as the process evolving until the end is reached. No trends of separation among roasting experiments performed under the same temperature level were instead observed, as shown in Fig. 2b, which shows the same PC1–PC2 score plot coloured according to the roasting temperature. Moreover, no differences according to the coffee species were found (data not shown). The corresponding loading plot (Fig. 2c) shows that, also in this case, many spectral regions contributed to explaining the spectral variability on the considered PCs. In fact, some peaks can be identified by looking at the loading profile of PC1, for example the ones between 8500 and 8000 cm^{-1}, at approximately 7000 cm^{-1} and in the range 5500–5000 cm^{-1}. All the variables belonging to these spectral regions pull the sample distribution towards positive PC1 score values. Therefore, these spectral regions can be related to the presence of chemical

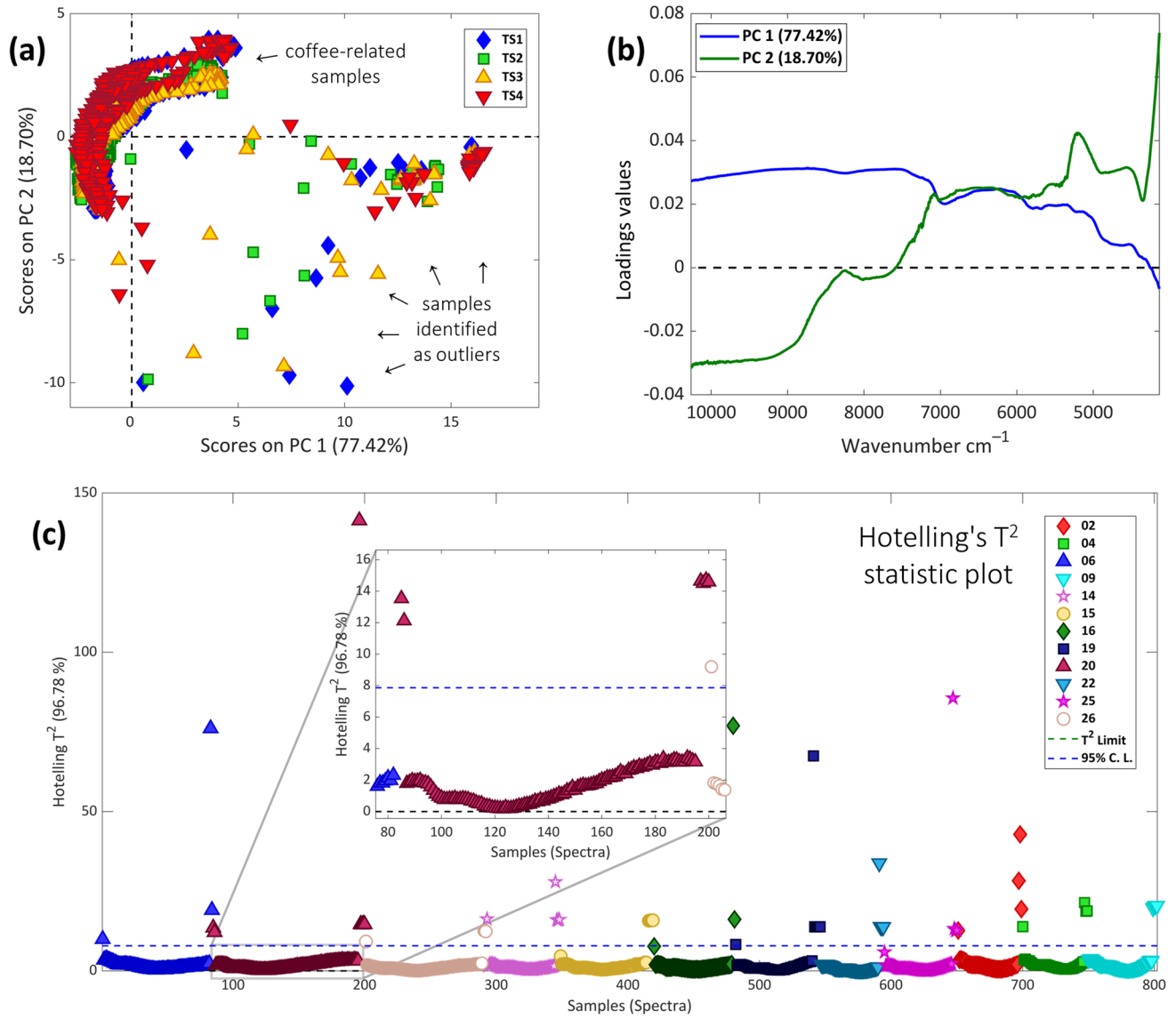


Fig. 1 PCA model results of the raw NIR spectra. **a** PC1 vs PC2 score plot coloured according to the roasting temperature level. **b** PC1 and PC2 loading plot. **c** From the same PCA model, the Hotelling $T.^2$ values of each spectrum collected during the experiments with Arabica coffee samples (plot coloured according to each individual single roasting experiment)

compounds whose NIR signal is more noticeable in the early stages of the roasting process. The spectral signals in the regions 7000 cm^{-1} and 5500–5000 cm^{-1} are mostly linked to the water content (first overtone and combination bands of O–H bonds stretching and bending in hydroxyl groups (Alessandrini et al. 2008; Santos et al. 2016b, a; Yergenson and Aston 2020b)), whereas the range 8500–8000 cm^{-1} is linked to C–H vibrations, associated with not only the content of organic acids, mainly chlorogenic and quinic acids (Esteban-Díez et al. 2004), but also physical features of the coffee beans, like dimension, shape and density (Catelani et al. 2017).

MCR-ALS Models

All MCR-ALS models well described the roasting process (explained variance > 99.9%, LOF < 0.63%, and standard deviation of the residuals lower than 0.0072).

Figure 3a reports the two spectral profiles (referred to as sopt1 and sopt2, respectively) corresponding to the two components identified by MCR-ALS at each roasting temperature level. All the spectral profiles obtained by the temperature-based MCR-ALS models show superimposable peaks and trends, demonstrating that the different tested roasting temperatures mainly bring the same physical and

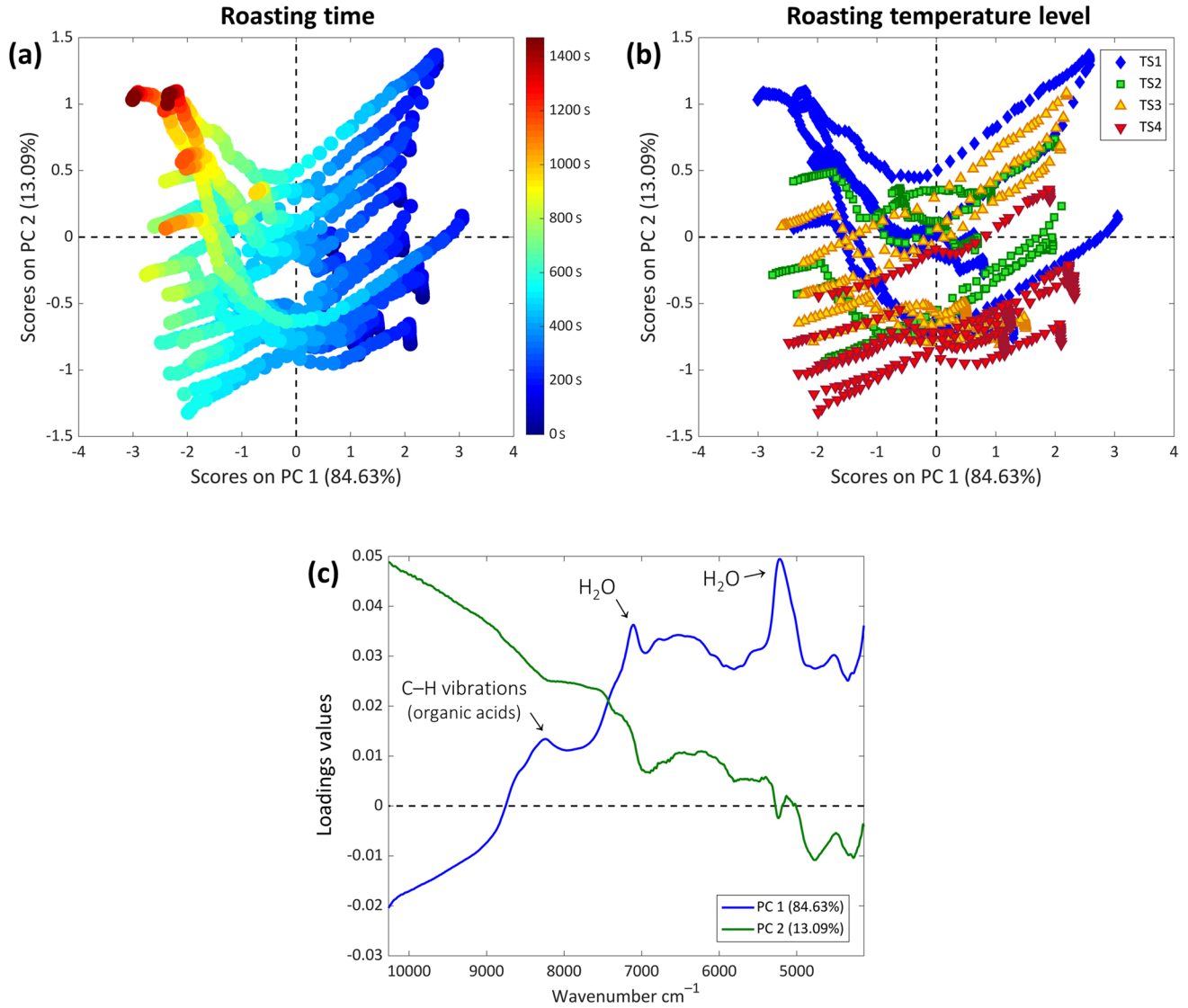


Fig. 2 PCA model results of the coffee NIR spectra. **a** PC1 vs PC2 score plot coloured according to roasting time (seconds). **b** The same PC1 vs PC2 score plot coloured according to the roasting temperature level. **c** PC1 and PC2 loading plot

chemical modifications to the coffee beans. Few differences were noticed for the extreme temperatures (TS1 and TS4).

Moreover, the comparison of the calculated profiles with the raw NIR spectra of a representative experiment (Fig. 3b) highlights the similarity, both in terms of vibrational pattern and shape, of the first and second spectral profiles to the initial and final spectra recorded during the roasting process, respectively. Such observation suggests that the MCR-ALS models identify two components corresponding to the typical spectra of the green and the roasted coffee beans.

Concerning the vibrational bands, both spectral profiles show a clear peak in the region 7500–6700 cm^{-1}, which is mostly related to the water content (6900 cm^{-1}, first overtone of the O–H stretching). The water signal is a broad band in the first spectral profiles, in agreement with the higher water content and the lower temperature at the beginning of the roasting process, whereas in the second profiles, representing the roasted beans, the band becomes sharper, and a shoulder appears at 6685 cm^{-1}.

In both profiles, within the region 6699–5400 cm^{-1}, essentially capturing the chemical information concerning C–H and S–H bonds (Catelani et al. 2017), two peaks at 5800 and 5700 are present, related to CH_2 first overtones present in long-chain fatty acids (Esteban-Díez et al. 2004).

The peak at 5200 cm^{-1} in the first spectral profiles, related to water content, is highly reduced in the second spectral profiles and it almost disappears in the one related to the trials performed at TS1. Indeed, the trials performed at TS1 reached very high weight losses (18.53% on average), comparable only to the experiments at TS4 (18.59%), being the roasting time 1500 s. The decrease of the

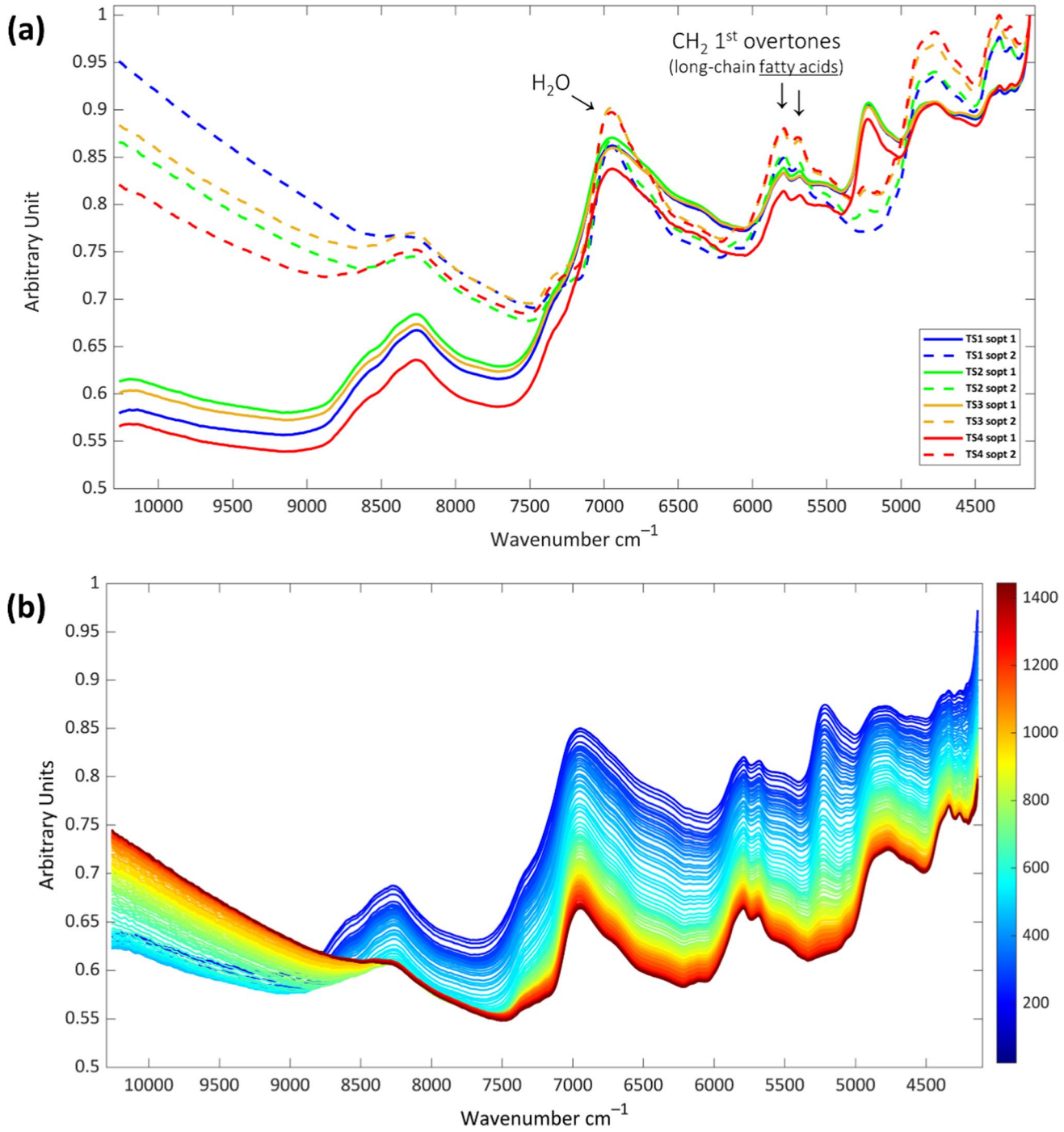


Fig. 3 MCR-ALS spectral profiles and FT-NIR spectra. **a** Spectral profiles (sopt1, sopt2) of the first and second component (solid and dashed line, respectively) computed by the MCR-ALS model at the different temperature levels: TS1 (blue), TS2 (green), TS3 (yellow), TS4 (red). **b** Example of a typical evolution of the NIR spectra of the coffee collected in real time from the beginning (blue curves) to the end (red curves) of the roasting process (experiment 17)

peak at 5200 cm^{-1} resulted in a visible small peak at 5170 cm^{-1} for the second spectral profile of the experiments performed at TS4.

In the region 4999–4000 cm^{-1}, containing chemical information related to C–H, N–H and C–O vibrations (Catelani et al. 2017), a broad band at 4810 cm^{-1} is present, assigned by Esteban-Díez et al. (Esteban-Díez et al. 2004) to carbohydrates. Additionally, two small peaks emerge in the first spectral profiles at 4343 cm^{-1} and at 4254 cm^{-1}, whose absorption increases according to roasting temperature in the roasted beans profiles, especially for the peak at 4343 cm^{-1}.

The appearance of these two peaks, attributable to combination modes between different vibrations of the C–H bond in CH_2 groups, was previously reported by several works (Alessandrini et al. 2008; Catelani et al. 2017) and may be related to the migration of oily compounds (mainly fatty acids) to the surface of the coffee beans, a phenomenon known to occur as a consequence of the formation of porosity during the roasting process (Schenker et al. 2000; Yergenson and Aston 2020b). The accumulation of fatty acids at the surface of the beans could have an impact also on the 6699–5400 cm^{-1} spectral range, whose absorbance indeed increases in the second profiles.

It is worth noticing that the reduced absorption of water after the roasting process influences the intensity of all the other vibrational bands, which are less covered by the water-related bands and resulted well resolved by the second MCR-ALS component. Moreover, the results confirmed what was observed by Gowen et al. (Gowen et al. 2013), i.e. that variations of the hydrogen-bonded network between water and other constituents have an impact on both components, as a multi-state system.

Beyond the chemical changes, a relevant modification in coffee bean roasting is also linked to changes in the beans' density. The second MCR-ALS profile is highly affected by coffee beans' physical changes, visible in the marked slope between 10,000 and 7300 cm^{-1} (Catelani et al. 2017). The increased background in this range also masks the actual variations that the vibrational bands between 8000 and 9000 cm^{-1}, assigned to the second overtone of C–H stretching of carbohydrates, lipids and amino acids, may undergo during the roasting process (Giraudo et al. 2019; Munyendo et al. 2021). As already observed in the raw spectra of experiment 17 (Fig. 3b), the curve slope increases with roasting time, indicating a change in the physical properties of the sample, which, in coffee roasting, could be ascribed to the change in density of the beans (Bustos-Vanegas et al. 2018). This behaviour was observed for all the considered batches.

The slope of the spectral profiles in this region is also linked to the roasting temperature. Indeed, the highest slope is visible for the batches roasted at TS1, whereas the lowest slope was recorded for the batches roasted at TS4. Intermediate temperatures (TS2 and TS3) are quite similar and are characterised by intermediate slopes. However, a correlation with the roasting time could also be hypothesized, as the TS1 and TS4 trials were the longest and shortest experiments, respectively.

The MCR-ALS concentration profiles of the first and second components (referred to as copt1 and copt2, respectively) relative to the whole spectral range (10,265–4135 cm^{-1}) are reported in Fig. 4 for the four considered temperature profiles. It is possible to notice that all the first profiles (filled markers) describe an inverse sigmoid: the first part of the process is characterised by almost constant concentration values extended for different times, according to the different temperature levels considered, then a decrease of the concentration leads the profiles to values tending towards an asymptote, except for the TS4 model. The second profiles (empty markers) are characterised by an opposite behaviour resulting in a sigmoidal trajectory for all the investigated temperatures, except for TS4 concentration profiles, which do not reach a plateau within the registered times. Two trials (TS2_R_2 and TS2_R_3; Fig. 4b) are characterised by a mild slope if compared to the steep one recorded for the other Robusta and Arabica trials performed at the same temperature. It is worth noting that the inspection of the raw spectra reveals that this difference can be attributed to the reduced spectral variation during these two roasting processes, and not to the variety characteristics.

Kinetic Fitting of the MCR-ALS Concentration Profiles

With the purpose of uncovering the kinetic critical points of the roasting process, the MCR-ALS concentration profiles were fitted with the sigmoid function described in the "Kinetic Fitting of the MCR-ALS Concentration Profiles" section, from which the maximum rate, acceleration and deceleration of the roasting phenomenon were calculated. The effects of temperature on the kinetic indexes calculated for the second concentration profiles were investigated by a two-way ANOVA also considering the coffee variety as a factor (A, Arabica, and B, Robusta). Table 1 reports the kinetic critical points obtained by the MCR-ALS models calculated considering both the whole spectral range (10,265–4135 cm^{-1}) and two reduced ranges (i.e. range 1, 6120–4961 cm^{-1}, and range 2, 9777–7417 cm^{-1}). When considering the whole spectral range, it is possible to notice that the lower temperature profile tested (TS1) resulted in a significant delay of the process in terms of maximum acceleration ($p < 0.01$), velocity ($p < 0.001$) and deceleration ($p < 0.001$), as expected. However, the increase in the temperature did not show any significant difference ($p > 0.05$) due to the variability of the different trials performed for each temperature level. The most relevant kinetic index to be considered is the maximum deceleration, as it can be useful to estimate the end of the roasting process, thus the point after which prolongation of the process could lead to an over-roasting of the coffee beans, which can negatively affect their sensory properties. The maximum deceleration of the second concentration profile for the batches roasted at TS1 occurred between 728 and 989 s, whereas for all the other tested temperatures it appeared between 534 and 679 s. No differences ($p > 0.05$) were instead ascribable to the variety factor,

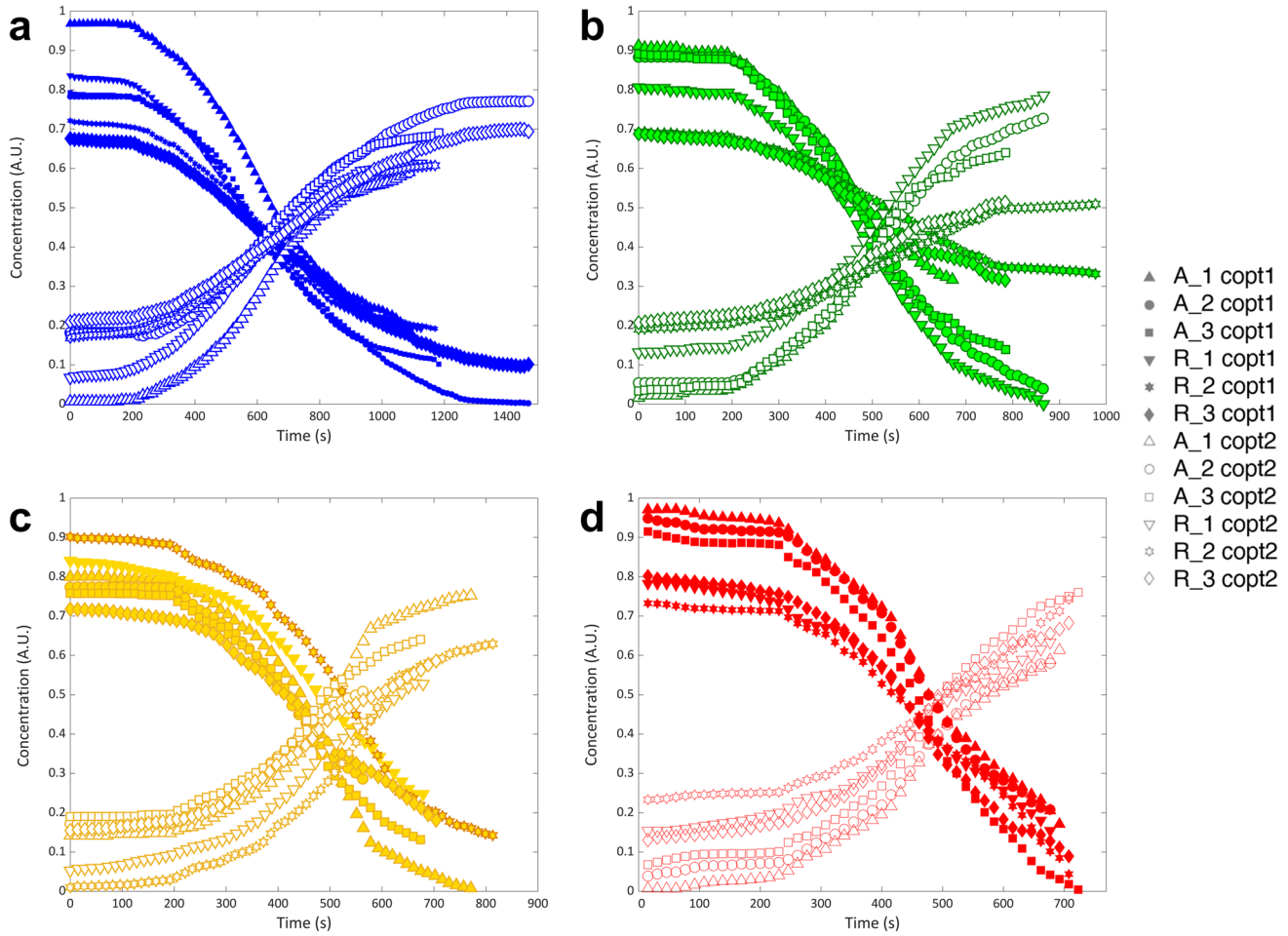


Fig. 4 MCR-ALS results, concentration profiles (copt) for trials at TS1 (**a**), TS2 (**b**), TS3 (**c**), TS4 (**d**). In all panels, full-coloured markers represent the trend of the first profile, edge-coloured markers represent the trend of the second profile. Different markers represent different trials: batch A_1 = black triangle; batch A_2 = black circle; batch A_3 = black squrea; batch R_1 = black inverted triangle; batch R_2 = asterisk; batch R_3 = diamond

indeed Arabica and Robusta responded in the same way to the roasting process for all the temperature levels tested.

When considering the concentration profiles of the MCR-ALS models obtained on the two reduced ranges, no statistically significant difference was uncovered for the maximum deceleration rate calculated for all the tested temperature profiles (Table 1, uppercase letters). However, the deceleration rates obtained by the two selected ranges differed from the ones obtained modelling the whole spectral range. In fact, as can be seen in Table 1, the kinetic indexes calculated for the whole spectral range showed to be higher than the ones referring to the reduced ranges. Nevertheless, the kinetics calculated for range 2 resulted in linearly correlated with the ones obtained by whole range modelling ($y = 0.895x$, $R^2 = 0.753$, $p < 0.01$).

To better visualise the relation among samples according to the kinetic parameters obtained from the MCR-ALS models, the computed kinetic values of each experimental dataset were organized into new data matrices composed of 24 samples and 3 variables (i.e. the kinetics parameters) and were inspected by PCA. The score plot obtained for the kinetic indexes calculated from the MCR-ALS model considering the whole spectral range (Fig. 5a) well represents the group of samples roasted at the lowest temperature (TS1), whereas samples roasted at higher temperatures resulted heterogeneously positioned in the left part of the plot. The score plot obtained considering the spectral range 1 (6120–4961 cm^{-1}; Fig. 5b) visually confirmed the lack of ability of this range to discriminate differences in the kinetic of the process according to temperature and variety. The second spectral range (9777–7417 cm^{-1}; Fig. 5c) seems to be the most affected by kinetic variations along the process when different operative conditions are considered: indeed, good groupings according to the roasting temperature were found. The increase in the temperature resulted in a decrease in PC1 values, even if some degrees of confusion can be noticed between samples

Table 1 Characteristic kinetic indexes calculated from the MCR-ALS concentration profiles describing the roasting behaviour models obtained considering the whole spectral range or the reduced ranges (range 1, 6120–4961 cm^{-1}; range 2, 9777–7417 cm^{-1})

	Whole spectral range (10,265–4135 cm^{-1})			Range 1 (6120–4961 cm^{-1})			Range 2 (9777–7417 cm^{-1})		
	Max acc	Max vel	Max dec	Max acc	Max vel	Max dec	Max acc	Max vel	Max dec
TS1_A_1	372^{b}	564^{b}	812^{bB}	192	467	826^{A}	413^{b}	564^{b}	729^{bA}
TS1_A_2	383^{b}	631^{b}	797^{bB}	260	467	686^{A}	343^{b}	562^{b}	783^{bA}
TS1_A_3	370^{b}	590^{b}	769^{bB}	287	495	742^{A}	356^{b}	535^{b}	701^{bA}
TS1_R_1	288^{b}	506^{b}	728^{bB}	151	356	536^{A}	370^{b}	522^{b}	714^{bA}
TS1_R_2	332^{b}	566^{b}	856^{bB}	193	427	594^{A}	413^{b}	594^{b}	787^{bA}
TS1_R_3	317^{b}	633^{b}	989^{bB}	207	439	633^{A}	303^{b}	578^{b}	851^{bA}
TS2_A_1	260^{a}	410^{a}	630^{a}	177	424	630	287^{b}	397^{a}	493^{a}
TS2_A_2	314^{a}	467^{a}	645^{a}	232	425	603	341^{b}	467^{a}	603^{a}
TS2_A_3	275^{a}	441^{a}	606^{a}	206	413	634	318^{b}	441^{a}	564^{a}
TS2_R_1	303^{a}	481^{a}	619^{a}	275	468	646	344^{b}	468^{a}	577^{a}
TS2_R_2	289^{a}	479^{a}	646^{a}	204	398	605	384^{b}	535^{a}	701^{a}
TS2_R_3	275^{a}	468^{a}	647^{a}	193	386	634	344^{b}	482^{a}	607^{a}
TS3_A_1	331^{a}	469^{a}	593^{bB}	248	442	688^{A}	345^{b}	455^{a}	578^{aA}
TS3_A_2	262^{a}	413^{a}	563^{bB}	235	440	563^{A}	290^{b}	372^{a}	494^{aA}
TS3_A_3	276^{a}	441^{a}	632^{bB}	262	523	674^{A}	304^{b}	399^{a}	482^{aA}
TS3_R_1	279^{a}	485^{a}	679^{bB}	225	430	609^{A}	293^{b}	403^{a}	526^{aA}
TS3_R_2	386^{a}	523^{a}	660^{bB}	221	427	674^{A}	372^{b}	496^{a}	619^{aA}
TS3_R_3	261^{a}	438^{a}	662^{bB}	219	411	662^{A}	288^{b}	397^{a}	523^{aA}
TS4_A_1	246^{a}	381^{a}	560^{bB}	300	436	587^{A}	260^{a}	340^{a}	436^{aA}
TS4_A_2	232^{a}	382^{a}	585^{bB}	259	423	598^{A}	232^{a}	355^{a}	518^{aA}
TS4_A_3	303^{a}	413^{a}	549^{bB}	207	385	631^{A}	234^{a}	344^{a}	427^{aA}
TS4_R_1	248^{a}	400^{a}	567^{bB}	205	400	595^{A}	205^{a}	332^{a}	457^{aA}
TS4_R_2	283^{a}	539^{a}	637^{bB}	283	567	637^{A}	241^{a}	339^{a}	427^{aA}
TS4_R_3	273^{a}	383^{a}	534^{bB}	205	356	562^{A}	218^{a}	370^{a}	562^{aA}

Legend of table's headings: *max acc*, maximum acceleration; *max vel*, maximum velocity; *max dec*, maximum deceleration. Lowercase letters (^{a}, ^{b}) in the table refer to differences within each column ($p < 0.05$), i.e. between temperatures and varieties; uppercase letters (^{A}, ^{B}) indicate differences between columns ($p < 0.05$), i.e. between the maximum deceleration times calculated for the considered spectral ranges. When no significant difference was observed, no apices are reported

roasted at TS2 and TS3. The visual inspection does not mark any systematic distributions of the samples according to the varieties.

The differences in describing the process kinetics according to all spectral ranges considered are undoubtedly related to the physicochemical information retained by each considered interval. Indeed, range 1 contains information concerning the major chemical changes involving C–H and S–H vibrations and C═O plus O–H and O–H vibrations (Catelani et al. 2018), whereas range 2 mainly considers physical changes, such as beans density, which contribute to the changes in slope from 9777 to 7300 cm^{-1}. Ideally, retaining both chemical and physical changes is advisable to describe the roasting progress; thus, it would be more advisable to consider the whole spectral range (10,265–4135 cm^{-1}) to follow the process in deep. However, in vision of simplified systems, range 2 (9777–7417 cm^{-1}) could be selected for the development of an ad hoc instrumentation. Indeed, the information retained in this region proved to be satisfactorily correlated with the information provided by the whole range considered and sensitive to process changes due to temperature modifications. Moreover, instruments working in this region could be developed with cheaper light sources and detectors, resulting in affordable sensors for the coffee industry.

Conclusions

The present study demonstrated the feasibility of a rapid, non-invasive, and automated method, based on NIR spectroscopy and multivariate data analysis, to perform real-time monitoring of the coffee roasting process.

The results, especially their visualization, demonstrate that PCA is a good diagnostic tool to identify and automatically remove NIR spectra not related to the coffee roasting

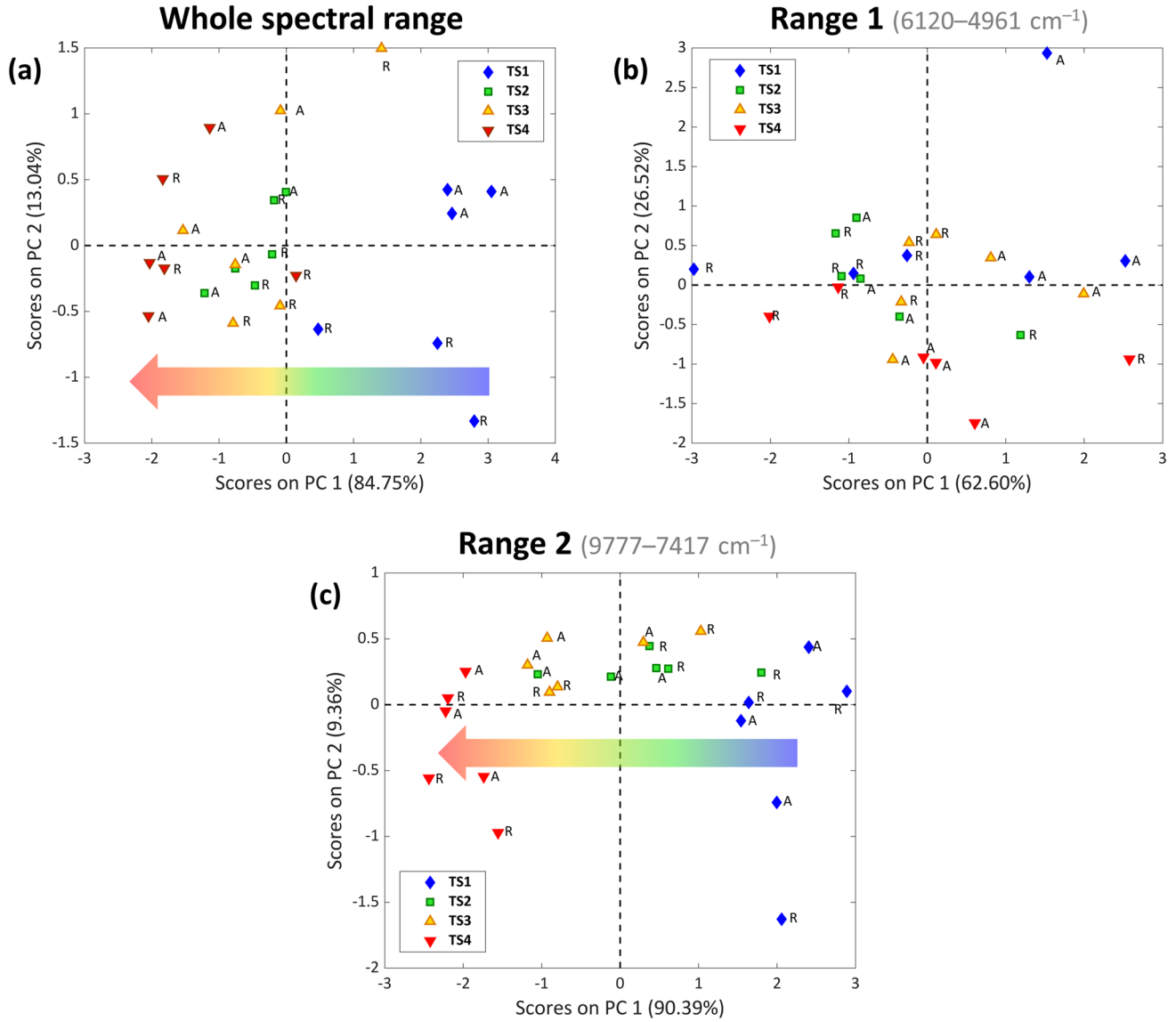


Fig. 5 PCA score plots of the kinetic parameters calculated from the MCR-ALS models. **a** Plot of the whole spectral range model. **b** Plot of the spectral range 1 model. **c** Plot of the spectral range 2 model. Letters "A" and "R" close to the samples' markers refer to Arabica and Robusta species, respectively

process (Fig. 1a, c) and, subsequently, to clearly describe the spectra (i.e. the time points) according to the roasting time (Fig. 2a). Further implementations of an automatic monitoring system could be equipped with a PCA step based on Hotelling's T^2 to filter out those spectra with strong experimental defects.

The application of MCR-ALS provided precious insights about the roasting dynamics under different process conditions: the spectral changes occurring during the roasting process could be efficiently modelled, and it is interesting to notice that these results were obtained with very simple models (only two components). The main changes that the coffee beans undergo during the roasting process could therefore be described from both the chemical and physical viewpoints, by interpreting the MCR-ALS resolved spectral profiles. Furthermore, the MCR concentration profiles were successfully fitted with a sigmoid function for estimating some noticeable kinetic steps, i.e. the maximum rate, acceleration, and deceleration of the coffee roasting process. From the interpretation of the time evolution of the roasting process, a significant ($p < 0.1$) delay was found for the roasting tests performed at the lowest temperature level with respect to those performed at the three higher temperature levels. In all cases, no significant differences were found based on the coffee varieties. A real-time implementation of an MCR-ALS-based system could envisage the use of both

historical data collected under the same conditions (roaster, experimental settings, coffee variety and quality) and "live" modelling of the evolving collected NIR information.

These considerations are coherent with the PAT paradigm, since the proposed approach would allow taking valuable process decisions in real-time and in a fully automatic way. Overall, the presented approach can be considered an important step towards the optimization of the coffee roasting process at the industrial scale since it allows taking valuable process decisions in real-time and in a fully automatic way. Moreover, in vision of simplified systems, a reduced spectral range (9777–7417 cm^{-1}) has proven promising for the development of simplified ad hoc instrumentation, potentially affordable also for smaller artisanal roasting plants.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12161-023-02473-w.

Author Contribution Silvia Grassi: conceptualization, methodology, software, validation, formal analysis, investigation, writing—original draft, writing—review and editing, visualization, data curation; Alessandro Giraudo: conceptualization, methodology, software, validation, formal analysis, investigation, writing—original draft, writing—review and editing, visualization, resources, data curation; Chiara Novara: investigation, writing—original draft, writing—review and editing, visualization, software, validation; Nicola Cavallini: validation, writing—review and editing, visualization; Francesco Geobaldo: project administration, funding acquisition, supervision; Ernestina Casiraghi: writing—review and editing, supervision; Francesco Savorani: project administration, funding acquisition, supervision, writing – review and editing.

Funding Open access funding provided by Politecnico di Torino within the CRUI-CARE Agreement.

Data Availability Data available on request to the authors.

Declarations

Competing interests The authors declare no competing interests.

Ethics Approval This study does not concern any studies with human participants or animals.

Consent to Participate Not applicable.

Conflict of Interest Silvia Grassi declares that she has no conflict of interest. Alessandro Giraudo declares that he has no conflict of interest. Chiara Novara declares that she has no conflict of interest. Nicola Cavallini declares that he has no conflict of interest. Francesco Geobaldo declares that he has no conflict of interest. Ernestina Casiraghi declares that she has no conflict of interest. Francesco Savorani declares that he has no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Alessandrini L, Romani S, Pinnavaia G, Rosa MD (2008) Near infrared spectroscopy: an analytical tool to predict coffee roasting degree. Anal Chim Acta 625:95–102. https://doi.org/10.1016/j.aca.2008.07.013

Andueza S, Vila MA, De Peña MP, Cid C (2007) Influence of coffee/water ratio on the final quality of espresso coffee. In: Journal of the Science of Food and Agriculture. John Wiley & Sons, Ltd, pp 586–592. https://doi.org/10.1002/jsfa.2720

Bertone E, Venturello A, Giraudo A et al (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

Bro R, De Jong S (1997) A fast non-negativity-constrained least squares algorithm. J Chemom 11:393–401. https://doi.org/10.1002/(SICI)1099-128X(199709/10)11:5%3c393::AID-CEM483%3e3.0.CO;2-L

Bro R, Smilde AK (2014) Principal component analysis. Anal. Methods 6:2812–2831. https://doi.org/10.1039/C3AY41907J

Bustos-Vanegas JD, Corrêa PC, Martins MA et al (2018) Developing predictive models for determining physical properties of coffee beans during the roasting process. Ind Crops Prod 112:839–845. https://doi.org/10.1016/j.indcrop.2017.12.015

Catelani TA, Páscoa RNMJ, Santos JR et al (2017) A non-invasive real-time methodology for the quantification of antioxidant properties in coffee during the roasting process based on near-infrared spectroscopy. Food Bioprocess Technol 10:630–638. https://doi.org/10.1007/s11947-016-1843-6

Catelani TA, Santos JR, Páscoa RNMJ et al (2018) Real-time monitoring of a coffee roasting process with near infrared spectroscopy using multivariate statistical analysis: a feasibility study. Talanta 179:292–299. https://doi.org/10.1016/j.talanta.2017.11.010

Czech H, Heide J, Ehlert S et al (2020) Smart online coffee roasting process control: modelling coffee roast degree and brew antioxidant capacity for real-time prediction by resonance-enhanced multi-photon ionization mass spectrometric (REMPI-TOFMS) monitoring of roast gases. Foods 9:627. https://doi.org/10.3390/foods9050627

de Juan A, Tauler R (2006) Multivariate curve resolution (MCR) from 2000: progress in concepts and applications. Crit Rev Anal Chem 36:163–176. https://doi.org/10.1080/10408340600970005

de Pires F, C, Pereira RGFA, Baqueta MR, et al (2021) Near-infrared spectroscopy and multivariate calibration as an alternative to the Agtron to predict roasting degrees in coffee beans and ground coffees. Food Chem 365:130471. https://doi.org/10.1016/J.FOODCHEM.2021.130471

De Luca S, De Filippis M, Bucci R et al (2016) Characterization of the effects of different roasting conditions on coffee samples of different geographical origins by HPLC-DAD, NIR and chemometrics. Microchem J 129:348–361. https://doi.org/10.1016/J.MICROC.2016.07.021

De Maria CAB, Trugo LC, Aquino Neto FR et al (1996) Composition of green coffee water-soluble fractions and identification of volatiles formed during roasting. Food Chem 55:203–207. https://doi.org/10.1016/0308-8146(95)00104-2

Dutra ER, Oliveira LS, Franca AS et al (2001) Preliminary study on the feasibility of using the composition of coffee roasting exhaust gas for the determination of the degree of roast. J Food Eng 47:241–246. https://doi.org/10.1016/S0260-8774(00)00116-3

Esteban-Díez I, González-Sáiz JM, Pizarro C (2004) Prediction of sensory properties of espresso from roasted coffee samples by near-infrared spectroscopy. Anal Chim Acta 525:171–182. https://doi.org/10.1016/j.aca.2004.08.057

Esteban-Díez I, González-Sáiz JM, Sáenz-González C, Pizarro C (2007) Coffee varietal differentiation based on near infrared spectroscopy. Talanta 71:221–229. https://doi.org/10.1016/j.talanta.2006.03.052

Franca AS, Mendonça JCF, Oliveira SD (2005) Composition of green and roasted coffees of different cup qualities. LWT - Food Sci Technol 38:709–715. https://doi.org/10.1016/j.lwt.2004.08.014

Franca AS, Oliveira LS, Oliveira RCS et al (2009) A preliminary evaluation of the effect of processing temperature on coffee roasting degree assessment. J Food Eng 92:345–352. https://doi.org/10.1016/j.jfoodeng.2008.12.012

Giraudo A, Grassi S, Savorani F et al (2019) Determination of the geographical origin of green coffee beans using NIR spectroscopy and multivariate data analysis. Food Control 99:137–145. https://doi.org/10.1016/j.foodcont.2018.12.033

Gowen AA, Amigo JM, Tsenkova R (2013) Characterisation of hydrogen bond perturbations in aqueous systems using aquaphotomics and multivariate curve resolution-alternating least squares. Anal Chim Acta 759:8–20. https://doi.org/10.1016/j.aca.2012.10.007

Grassi S, Alamprese C (2018) Advances in NIR spectroscopy applied to process analytical technology in food industries. Curr Opin Food Sci 22:17–21. https://doi.org/10.1016/J.COFS.2017.12.008

Grassi S, Alamprese C, Bono V et al (2013) Monitoring of lactic acid fermentation process using Fourier transform near infrared spectroscopy. J near Infrared Spectrosc 21:417–425. https://doi.org/10.1255/jnirs.1058

Heide J, Czech H, Ehlert S et al (2020) Toward smart online coffee roasting process control: feasibility of real-time prediction of coffee roast degree and brew antioxidant capacity by single-photon ionization mass spectrometric monitoring of roast gases. J Agric Food Chem 68:4752–4759. https://doi.org/10.1021/acs.jafc.9b06502

Iaccheri E, Laghi L, Cevoli C et al (2015) Different analytical approaches for the study of water features in green and roasted coffee beans. J Food Eng 146:28–35. https://doi.org/10.1016/j.jfoodeng.2014.08.016

Li Vigni M, Durante C, Cocchi M (2013) Exploratory data analysis. In: Data handling in science and technology. Elsevier Ltd, pp 55–126. https://doi.org/10.1016/B978-0-444-59528-7.00003-X

Munyendo L, Njoroge D, Hitzmann B (2021) The potential of spectroscopic techniques in coffee analysis—a review. Process 2022, Vol 10, Page 71 10:71 https://doi.org/10.3390/PR10010071

Pizarro C, Esteban-Díez I, González-Sáiz JM, Forina M (2007) Use of near-infrared spectroscopy and feature selection techniques for predicting the caffeine content and roasting color in roasted coffees. J Agric Food Chem 55:7477–7488. https://doi.org/10.1021/jf071139x

Santos JR, Lopo M, Rangel AOSS, Lopes JA (2016a) Exploiting near infrared spectroscopy as an analytical tool for on-line monitoring of acidity during coffee roasting. Food Control 60:408–415. https://doi.org/10.1016/j.foodcont.2015.08.007

Santos JR, Viegas O, Páscoa RNMJ et al (2016b) In-line monitoring of the coffee roasting process with near infrared spectroscopy: measurement of sucrose and colour. Food Chem 208:103–110. https://doi.org/10.1016/j.foodchem.2016.03.114

Schenker S, Handschin S, Frey B et al (2000) Pore structure of coffee beans affected by roasting conditions. J Food Sci 65:452–457. https://doi.org/10.1111/j.1365-2621.2000.tb16026.x

Shan J, Suzuki T, Ogawa Y, Kondo N (2015) Coffee roasting degrees prediction in terms of weight loss with selected wavebands based on near-infrared spectroscopy. Eng Agric Environ Food 8:195–199. https://doi.org/10.1016/j.eaef.2015.10.003

Wall ME, Rechtsteiner A, Rocha LM (2005) Singular value decomposition and principal component analysis. In: A practical approach to microarray data analysis. Kluwer Academic Publishers,pp 91–109

Yergenson N, Aston DE (2020b) Monitoring coffee roasting cracks and predicting with in situ near-infrared spectroscopy. J Food Process Eng 43 https://doi.org/10.1111/jfpe.13305

Yergenson N, Aston DE (2020a) Online determination of coffee roast degree toward controlling acidity. J near Infrared Spectrosc 28:175–185. https://doi.org/10.1177/0967033520924493

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.