Food Research International 176 (2024) 113800

Contents lists available at ScienceDirect

Food Research International

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

Evaluation of spontaneous fermentation impact on the physicochemical properties and sensory profile of green and roasted arabica coffee by digital technologies

Hanjing Wu [a], Claudia Gonzalez Viejo [b,*], Sigfredo Fuentes [b,d], Frank R. Dunshea [b,c], Hafiz A. R. Suleria [a,*]

[a] School of Agriculture, Food and Ecosystem Sciences, Faculty of Sciences, The University of Melbourne, Parkville 3010, VIC, Australia
[b] Digital Agriculture, Food and Wine Group, School of Agriculture, Food and Ecosystem Sciences, Faculty of Sciences, The University of Melbourne, Parkville 3010, VIC, Australia
[c] Faculty of Biological Sciences, The University of Leeds, Leeds, UK
[d] Tecnologico de Monterrey, School of Engineering and Sciences, Ave. Eugenio Garza Sada 2501, Monterrey, N.L. 64849, Mexico

ARTICLE INFO

Keywords:
Fermented coffee
Fermentation
E-nose
Sensory analysis
Aromatic volatile compounds
Machine learning

ABSTRACT

There is a growing demand for specialty coffee with more pleasant and uniform sensory perception. Wet fermentation could modulate and confer additional aroma notes to final roasted coffee brew. This study aimed to assess differences in volatile compounds and the intensities of sensory descriptors between unfermented and spontaneously fermented coffee using digital technologies. Fermented (F) and unfermented (UF) coffee samples, harvested from two Australia local farms Mountain Top Estate (T) and Kahawa Estate (K), with four roasting levels (green, light-, medium-, and dark-) were analysed using near-infrared spectrometry (NIR), and a low-cost electronic nose (e-nose) along with some ground truth measurements such as headspace/gas chromatography–mass spectrometry (HS-SPME-GC-MS), and quantitative descriptive analysis (QDA ®). Regression machine learning (ML) modelling based on artificial neural networks (ANN) was conducted to predict volatile aromatic compounds and intensity of sensory descriptors using NIR and e-nose data as inputs. Green fermented coffee had significant perception of hay aroma and flavor. Roasted fermented coffee had higher intensities of coffee liquid color, crema height and color, aftertaste, aroma and flavor of dark chocolate and roasted, and butter flavor ($p < 0.05$). According to GC-MS detection, volatile aromatic compounds, including methylpyrazine, 2-ethyl-5-methylpyrazine, and 2-ethyl-6-methylpyrazine, were observed to discriminate fermented and unfermented roasted coffee. The four ML models developed using the NIR absorbance values and e-nose measurements as inputs were highly accurate in predicting (i) the peak area of volatile aromatic compounds (Model 1, R = 0.98; Model 3, R = 0.87) and (ii) intensities of sensory descriptors (Model 2 and Model 4; R = 0.91), respectively. The proposed efficient, reliable, and affordable method may potentially be used in the coffee industry and smallholders in the differentiation and development of specialty coffee, as well as in process monitoring and sensory quality assurance.

1. Introduction

Nowadays, coffee consumers are paying more attention to the rewarding experience of coffee consumption, which requires coffee to offer more nutritional benefits and pleasant sensory quality (de Melo Pereira et al., 2019). The sensory properties of end-products could be attributed to many variables, including physical characteristics, chemical composition, and post-harvest primary and secondary processing of coffee beans (Sualeh et al., 2020). Spontaneous fermentation is one of the primary postharvest processing methods, where a wide range of indigenous microbiota perform microbial activities on the de-pulped beans for 6 to 72 h (Masoud et al., 2004; Scholz et al., 2019).

Involved microorganisms during fermentation would utilize the carbohydrates, proteins, and phenolic compounds, especially reducing sugar, as carbon sources for growth, which could largely determine the remaining content of free sugars and amino acids. Moreover, the

* Corresponding authors.
E-mail addresses: cgonzalez2@unimelb.edu.au (C. Gonzalez Viejo), hafiz.suleria@unimelb.edu.au (H.A.R. Suleria).

https://doi.org/10.1016/j.foodres.2023.113800
Received 20 September 2023; Received in revised form 24 November 2023; Accepted 2 December 2023
Available online 9 December 2023

generation of microbial metabolites, such as organic acids, alcohols, esters, ketones, and aldehydes, may migrate into the coffee beans during long-time fermentation (Elhalis et al., 2022). Most of them are essential aroma precursors and play a role in the production of volatile aromatic compounds and the development of coffee flavor during subsequent drying and roasting (da Silva Oliveira et al., 2022). Pereira et al. (2022) reported that spontaneously fermented coffee had a higher level of trigonelline, which is responsible for the coffee bitterness and astringency, and the generation of furans correlated to sweet, caramel, and burnt aromas (Bressani et al., 2021). Thus, fermentation is believed to modulate or confer additional aroma notes to the final coffee beverage and influence the sensory qualities both directly and indirectly.

Generally, the determination of non-volatile and volatile compounds in coffee and the evaluation of their sensory properties are achieved through liquid chromatography (LC), gas chromatography (GC), and quantitative descriptive analysis (QDA®) with a panel of Q-graders, which require high cost of time, money, and learning. Furthermore, considering the recognition threshold, the sensory analysis could induce stimulus errors and become less objective and senseless. Therefore, developing novel, rapid, and reliable analyses, such as near-infrared (NIR) and electronic noses (e-noses), is beneficial as an alternative to advanced but expensive analytical methods. NIR and e-nose coupled with artificial neural networks (ANNs) have been explored for several purposes, such as coffee varietal differentiation (Buratti et al., 2015), determination of coffee geographic origin and postharvest processing type (Giraudo et al., 2019; Wu et al., 2023), and the prediction of coffee roasting degree (Bertone et al., 2016; Gonzalez Viejo et al., 2021). However, individual usage of NIR and e-nose also has limitations when analyzing trace composition in complex samples (Wang, 2019).

This study aimed to evaluate the impacts of spontaneous fermentation on the physicochemical properties, the pattern of volatile aromatic compounds, and the sensory perception of Arabica coffee. For this, NIR, e-nose, GC-MS, and QDA tests with machine learning modelling were used to evaluate, determine, perceive, and predict the differences between unfermented and fermented coffee and principal components analysis (PCA) was used to assess the grouping relationship between variables and sensory profile of coffee. The interactions among postharvest processing, chemical composition, and sensory quality could be better understood and further investigated to optimize coffee processing.

2. Materials and methods

2.1. Sample preparation

There was a total of four bags (500 g each bag) (n = 4), two bags of unfermented (n = 2) and two fermented (n = 2) Kenyan arabica green coffee beans. One of each type was planted and provided by two separated local coffee farms in Australia: (1) Mountain Top Estate and (2) Kahawa Estate. Green coffee beans were respectively roasted into light roast level (196 °C, 10 mins), medium roast level (210 °C, 10 mins), and dark roast level (225 °C, 10mins) using a coffee roaster (Cafemasy, model SCR-301, Guangzhou, China). Coffee roaster was pre-heated under relevant temperature for 15 min before actual roasting coffee beans to ensure the achievement of required temperature. Green and roasted coffee beans were ground into a powder with a mean particle size by a coffee grinder (Breville Smart Grinder™ Pro, model BCG820BSSXL, Melbourne, VIC, Australia) after cooling down to ambient temperature (25 °C) with natural air flush (Soocas Hair dryer, model H3, Shenzhen, Guangdong, China). Interval breaks were applied every five grinding operations to avoid the variation from machine heating. Ground coffee (2.5 g) was brewed into liquid using the Breville Creatista® Plus espresso machine (Breville Pty Ltd., Sydney, New South Wales, Australia) under Expresso mode of a constant pouring volume of 30 mL at 78 °C. Therefore, 16 coffee liquid brewed from different processed beans (Table 1) were used for the following analysis.

Table 1
Coffee samples involved in this study, including the origin farm, postharvest processing type, roasting degree, and abbreviation.

Processing type	Origin farm	Green (unroasted) (G)	Light roasted (L)	Medium roasted (M)	Dark roasted (D)
Unfermented (UF)	Kahawa Estate (K)	UFKG	UFKL	UFKM	UFKD
	Mountain Top Coffee (T)	UFTG	UFTL	UFTM	UFTD
Fermented (F)	Kahawa Estate (K)	FKG	FKL	FKM	FKD
	Mountain Top Coffee (T)	FTG	FTL	FTM	FTD

2.2. Total soluble solids and color measurement

Total soluble solids in coffee brew and replicates were measured in degrees Brix (°Brix) using an optical refractometer Alla France REFBX010 (Alla France Sarl, ChemilléMelay, France) with a range of measurement of 0–32 Brix. All coffee brews were measured at 40 °C, and the refractometer was rinsed with distilled water and dried between measurements to avoid cross-contamination. The color of all coffee samples brewed from different roasted beans was measured in triplicates using a handheld colorimeter NIX (Nix Pro Colour Sensor™, Nix Sensor Ltd., Ontario, CA). The CieLab color coordinates (L*, a*, and b* colorimetric, unitless) were recorded and analyzed.

2.3. Near-Infrared spectroscopy (NIR) analysis

A portable NIR device microPHAZIR™ (RX Analyzer; Thermo Fisher Scientific, Waltham, MA, USA) was used to evaluate the coffee brew according to the method described by Gonzalez Viejo et al. (2018a) with the spectral range of 1596 to 2396 nm every 7–9 nm in triplicates and three measurements per replicate. The absorbance was measured using a Tungsten light bulb with the measurement time less than three seconds. A Whatman® filter paper (Whatman plc. Maidstone, UK; qualitative grade 3, 7.0 cm) was submerged in each coffee sample to be measured. A white background was placed at the top to avoid environmental signal noise during the measuring. The absorbance values from the dry and empty filter paper were subtracted from the wet filter paper with the samples to obtain only the chemical fingerprinting of the coffee sample, as described by Gonzalez Viejo et al. (2018a). The first derivative of NIR data was obtained using Savitzky Golay filters in The Unscrambler X ver.10.3 software (CAMO Software, Oslo, Norway).

2.4. Electronic nose (e-nose) description and data extraction

A low-cost and portable e-nose developed by the Digital Agriculture Food and Wine Group from the University of Melbourne (DAFW-UoM) was used to assess the coffee samples, according to Gonzalez Viejo et al. (2020). The device was composed of nine different gas sensors: (i) MQ3: Alcohol (ethanol), (ii) MQ4: Methane (CH_4), (iii) MQ7: Carbon monoxide (CO), (iv) MQ8: Hydrogen (H_2), (v) MQ135: NH_3/alcohol/benzene, (vi) MQ136: hydrogen sulfide (H_2S), (vii) MQ137: NH_3, (viii) MQ138: benzene/alcohol/NH_3 and (ix) MG811: Carbon dioxide (CO_2). Each brewed coffee sample in beaker was immediately exposed to sensors placed on the top of beaker for 1 min after 30 s baseline readings. The outputs were analyzed using a customized Matlab® R2021a code (Mathworks, Inc., Natick, MA, USA) developed by the DAFW-UoM, which displays the curves to select the most stable area of the signals and subdivide it into 10 sections to calculate 10 mean values per curve automatically for further analysis.

2.5. Identification and quantification of volatiles by HS-SPME-GC-MS

Volatile compounds in coffee samples were analyzed by HS-SPME-GC-MS based on the methods described by Gonzalez Viejo et al. (2021) and Wu et al. (2022) with some modifications. GC-MS analysis was conducted via a gas chromatograph (6850 series II Network GC System, Agilent Technologies, USA) coupled to a headspace solid-phase microextraction (HS-SPME) system (PAL RSI I20, Switzerland) and a mass spectrometer (5973Network Mass Selective Detector, Agilent Technologies, USA).

A 30 m DB-Wax capillary column (Agilent Technologies, USA) with 0.25 mm internal diameter and 0.25 μm film thickness was fitted. The carrier gas was helium with 60 kPa column head pressure. In headspace sampling system, each sample was incubated for 15 mins at 40 °C and then 15 mins extraction and being sampled by 65 μm PDMS/DVB fiber (Fused Silica, Sigma Aldrich, USA). The compounds were desorbed for 6 mins in splitless mode.The GC oven program was modified and set as follows: 40 °C for 5 min followed by an increase to 190 °C with the rate of 5 °C/min for 5 min; subsequently, the temperature reached 230 °C at a rate of 10 °C/min and maintained for 7 min. The acquisition was set to SCAN mode (35–350 m/z). The solvent delay time was 2 min.

4-Octanol (20 μL 100 mg/L) was added into vials as internal standard and mixed and vortexed with 2.5 g of coffee sample. The linear retention index (LRI) of detected volatile compounds with >85% certainty was calculated using the alkane standard (C_7–C_{30}). Calculated LRI and mass spectrum of detected volatile compounds were compared to the data in the NIST Chemistry WebBook spectrum library (NIST2017) and NIST mass-spectra database, respectively. Semi-quantification was conducted by comparing the response area of the target compound and a closely eluted compound with known concentration after LRI and compound MS were confirmed.

2.6. Descriptive sensory evaluation

Twelve participants from the University of Melbourne (UoM; Ethics ID:1953926) were pre-screened as regular coffee drinkers and recruited into the sensory panel. All the panelists were professionally trained to detect basic tastes and aromas using International Standard methodology (ISO 8586-1: 1993E Sensory analysis – General guidelines for the selection, training and monitoring of selected assessors and expert sensory assessors, and quality control procedures) (ISO 1993) (Gonzalez Viejo et al., 2018b). This training was conducted in five sessions of 60 mins each. As suggested in the above International Standard, briefly, sodium chloride (0.5 g/L), sucrose (16 g/L), caffeine (0.5 g/L), monosodium glutamate (0.6 g/L), and citric acid (1 g/L) were used for the general training of basic tastes, salty, sweet, bitter, umami, and sour, respectively. Subsequently, aroma references (Le Nez du Vin ®: The Master kit) (Le Nez du Vin, France), including straw, clove, vanilla, honey, butter, leather, toast, coffee blossom, coffee pulp, maple syrup, and dark chocolate, were used to train the familiarization with relevant aromas generally found in coffee.

The selection of descriptors for the test was carried out using the quantitative descriptive analysis (QDA®) method in blind tasting sessions using the different coffee samples to generate consensus and agreement on a set of the most relevant attributes. For QDA®, the training sessions and selection of descriptors for arabica coffee samples consisted of nine sessions of 60 mins each and divided as follows: (i) two sessions for fermented green coffee, (ii) two sessions for fermented light-roasted coffee, (iii) two sessions for fermented medium-roasted coffee, (iv) two sessions for fermented dark-roasted coffee, (v) two sessions for unfermented green coffee, (vi) two sessions for unfermented light-roasted coffee, (vii) two sessions for unfermented medium-roasted coffee, (viii) two sessions for unfermented dark-roasted coffee, (ix) one session for a mix of all types of coffee. To assess the panel performance during the training, a combination of cluster analysis, standard deviation, ANVOA, and spider chart (data not shown) were developed to assess significant differences within the panelists for each descriptor.

The single double-blind sensory session was conducted to evaluate the intensity of sensory descriptors for the 16 coffee samples in the focus group-type room in the sensory laboratory (School of Agriculture, Food and Ecosystem Sciences, UoM). Sample coffee brews (30 mL) were served at 70 °C in white paper shot cups labeled with 3-digit random codes. Water and plain water crackers were offered as palate cleansers. A total of 28 sensory descriptors with the rating of their intensity in a 15 cm non-structured scale in the questionnaire (Table 2) were displayed on RedJade® (RedJade Sensory Solutions, LLC, Martinez, CA, USA).

2.7. Statistical analysis and machine learning (ML) modelling

GraphPad Prism 9 (GraphPad Prism version 9.0 for Windows, GraphPad Software, La Jolla, California, USA) was used for data visualization. One-way and two-way analysis of variance (ANOVA), along with Fisher's least significant difference (LSD) as *post hoc* test ($\alpha = 0.05$), were performed via XLSTAT 2020.3.1 (Addinsoft, New York, USA) to assess significant differences among samples, processing methods, and the interactions of fermentation and roasting. Furthermore, a multivariate data analysis based on principal components analysis (PCA) was performed based on covariance to find relationships between variables and their associations with the samples.

Four ML models were constructed using a Matlab® R2019b code developed by the Digital Agriculture Food and Wine group from The

Table 2
Sensory descriptors evaluated and anchors used during coffee sensory evaluation.

Descriptor	Anchors	Descriptor	Anchors
Appearance			
Liquid Turbidity *(clarity)*	Clear - Muddy	Crema Height	Absent - High
Liquid Color	Light - Dark	Crema Color	Light - Dark
Aroma ***(as an odor, sensed through the nose and retronasal olfaction)***			
Aroma Floral	Absent - Intense	Aroma Dark Chocolate	Absent - Intense
Aroma Hay	Absent - Intense	Aroma Sweet	Absent - Intense
Aroma Butter	Absent - Intense	Aroma Spices	Absent - Intense
Aroma Earthy	Absent - Intense	Aroma Roasted	Absent - Intense
Taste ***(the sense experienced by the tongue)***			
Sweetness	Absent - Intense	Sourness	Absent - Intense
Bitterness	Absent - Intense	Astringency *(a tactile taste felt as a dry, rough feeling)*	Absent - Intense
Flavor ***(a combination of both aroma and taste)***			
Flavor Floral	Absent - Intense	Flavor Dark Chocolate	Absent - Intense
Flavor Hay	Absent - Intense	Flavor Sweet	Absent - Intense
Flavor Butter	Absent - Intense	Flavor Spices	Absent - Intense
Flavor Earthy	Absent - Intense	Flavor Roasted	Absent - Intense
Smoothness *(no overintense tastes of sour/bitter/ astringent)*	Absent - Intense	Body *(a tactile sense of density/viscosity)*	Light - Full
Aftertaste *(the tastes and aromas left in the mouth after swallowing)*	0 Second - >5 Seconds	Overall Quality *(overall sensory performance)*	Unacceptable - Extraordinary

University of Melbourne to assess 17 training algorithms and find the most accurate models with no under- or overfitting in a loop (Gonzalez Viejo et al., 2019). Models 1 and 2 (Fig. 1**a**) were developed using regression artificial neural networks (ANN) with the Levenberg-Marquardt algorithm to predict the peak area of 26 volatile aromatic compounds using the absorbance values of NIR (1596–2396 nm; Model 1) and e-nose results (Model 2) as inputs. Samples were divided as 70% for the training stage, 15% for testing, and 15% for validation using a performance algorithm based on mean squared error (MSE).

Models 3 and 4 (Fig. 1**b**) were developed using regression ANN with the data of NIR and e-nose as inputs, respectively, to predict 28 sensory descriptors. Both models were developed using the Bayesian Regularization algorithm. Data were randomly divided into 70:30 for training and testing, respectively, using a performance algorithm based on MSE. A neuron number trimming exercise was conducted with 3, 5, 7, and 10 neurons to assess the best performance and ensure no under or overfitting based on accuracy and MSE values.

3. Results and discussion

3.1. Descriptive sensory evaluation

Fig. 2 shows significant differences ($p < 0.05$) among all coffee samples for the 16 sensory descriptors for unfermented (Fig. 2**a**) and fermented (Fig. 2**b**) coffee T and K (T = Mountain Top Coffee; K = Kahawa Estate, according to Table 1), except for sweetness. FTD had the highest intensities for liquid color (14.31), crema height (thickness) (11.66), crema color (8.50), dark chocolate flavor (5.63), and aftertaste (10.89). However, UFTD had a relatively higher intensity for bitterness (9.04). When considering the processing method as qualitative variables, significant differences ($p < 0.05$) (**Table A1**) only existed in the liquid color and bitterness among coffee samples. Overall, fermented coffee had lighter liquid color and lower perceived bitterness. As for the interactions of fermentation and roasting intensity as qualitative variables, significant differences ($p < 0.05$) (**Table A2**) could be observed among coffee samples for the same 15 sensory descriptors but still excluding sweetness.

Overall, it could be deducted that spontaneous fermentation could not be a determining factor for the sensory perception of coffee brew, whereas roasting seems to play a pivotal role. This observation is consistent with the findings reported by Elhalis et al. (2021) and da Silva Oliveira et al. (2022). De Bruyn et al. (2017) argued that the microbiota of the planting region and processing methods of coffee beans could contribute to the sensory quality of coffee drinks. Pereira et al. (2022) conducted a sensory evaluation on naturally processed and self-fermented green coffee brewed from the coffee beans harvested from the farm in Brazil. Fermented coffee was found to have a relatively higher dominance of woody, herbaceous, and fruity attributes. In accordance with this study, green fermented coffee had higher intensities for hay aroma (8.313) and flavor (5.197).

Dark-roasted fermented coffee had a relatively higher intensity for liquid color (13.05) compared to the unfermented. Nursten (2005) and Elhalis et al. (2022) commented that fermented coffee beans contain higher amounts of aspartic and glutamic acids, which are intermediate browning-producing amino acids during Maillard reaction. Therefore, dark-roasted fermented coffee brew was perceived as a darker liquid color.

Sweetness and sourness are commonly negatively correlated with each other at medium and high concentration but variably at low concentration (da Silva Oliveira et al., 2022). The perceived intensity of sweetness and sourness in a mixture could be less at the same concentration level (Zamora et al., 2006). Although fermented coffee brew was assessed with significantly higher intense sour taste (FTD: 6.26; FTM: 7.30; UFTD: 4.37; UFTM: 6.28), non-significant variance in sweetness among coffee samples could probably because of both low concentration and mixture suppression.

The weaker perception of bitterness from the fermented coffee sample in this study could be attributed to the occurrence of the degradation of caffeine and protein and the loss of phenolics (chlorogenic acid and quinic acid) during fermentation (da Silva Oliveira et al., 2022; Elhalis et al., 2022; Fujimoto et al., 2021; Lee et al., 2015). Bressani et al. (2021) observed consistent results from fermented coffee after filamentous fungi and bacteria fermentation, where caffeine could be degraded into xanthine or dimethylxanthine. Moreover, after fermentation, more free amino acids released from beans were observed by Robinson (2014), which indirectly presented the severe protein degradation occurring inside the coffee beans. Thus, the perceived strength of bitterness from fermented coffee was significantly lower.

3.2. Multivariate data analysis

Fig. 3 shows the PCA of the e-nose measurements, 21 aromatic

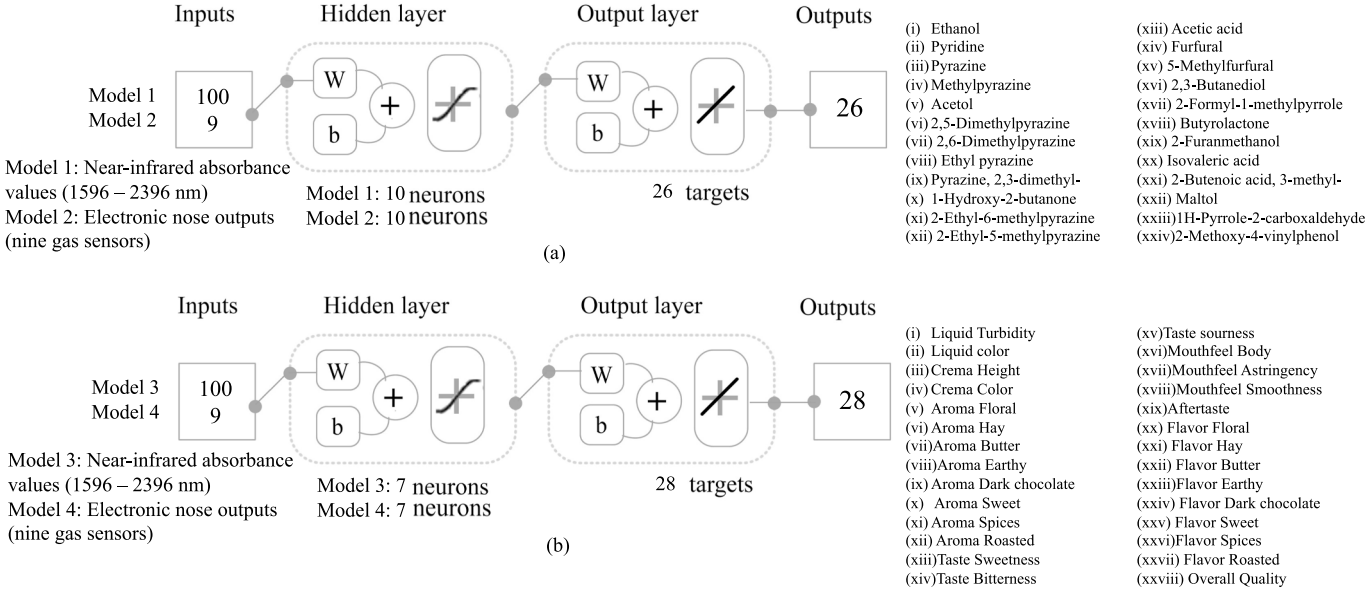


Fig. 1. Diagrams of machine learning including the inputs, targets, and the number of neurons used for four regression models: Models 1 and 2 (a), 3 and 4 (b).

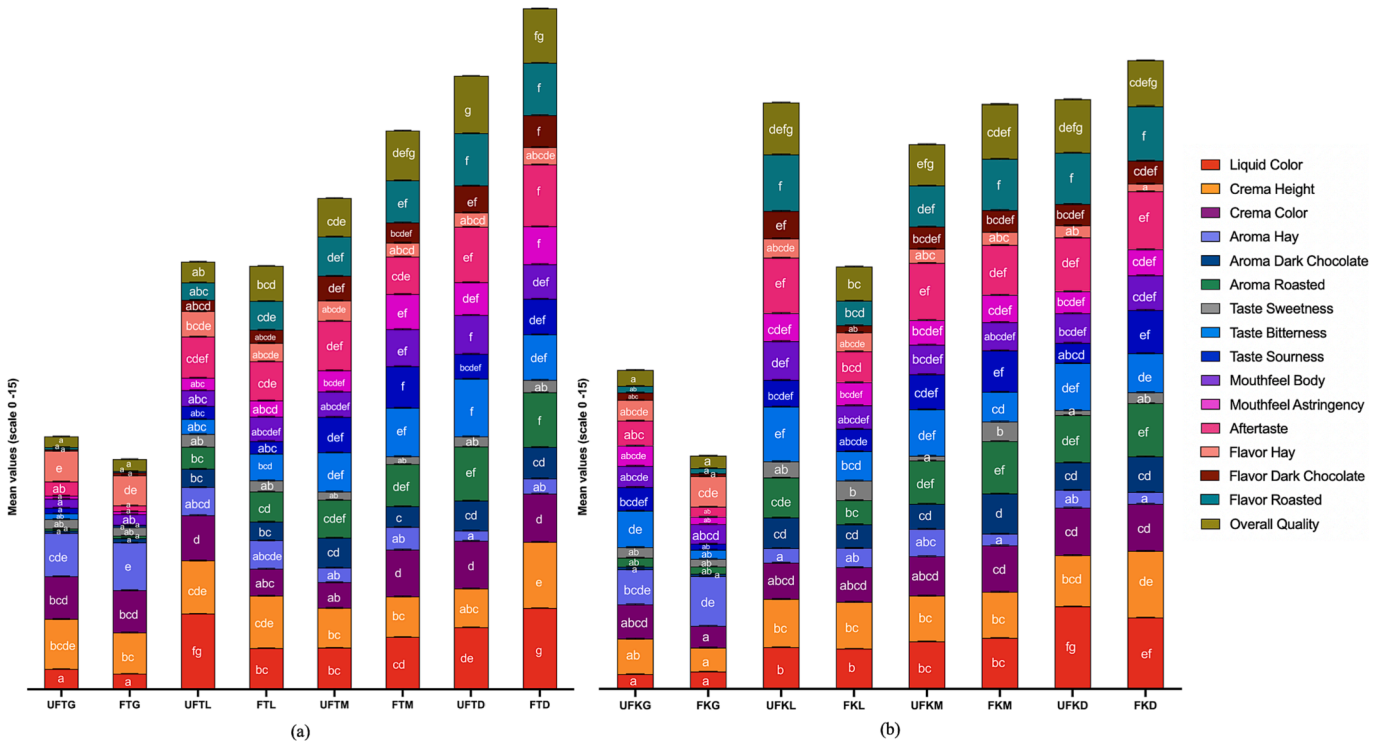


Fig. 2. Mean evaluation values of 16 out of 28 sensory descriptors showed significant differences for unfermented (UF) and fermented (F) coffee brewed from coffee beans T (a) and K (b). Different lowercase letters (a-g) depict significant differences between samples based on one-way ANOVA and Fisher's least significant difference (LSD) *post hoc* test ($\alpha < 0.05$). The sample abbreviations are listed in Table 1. Standard error ranges from 0.06 to 0.11.

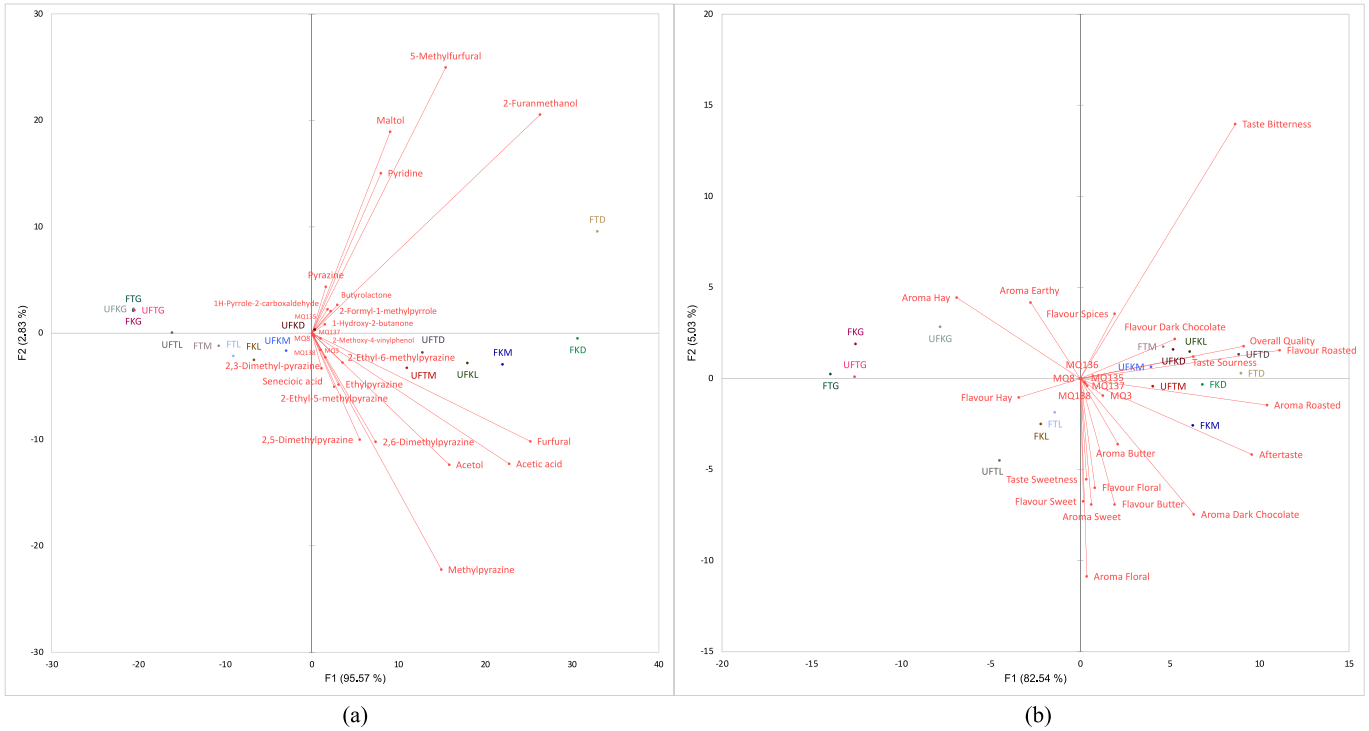


Fig. 3. PCA from e-nose detection and 21 aromatic volatile compounds (a), and 19 out of 28 sensory descriptors relevant to aroma, flavor, and taste (b). The sample abbreviations are listed in Table 1. X-axis represents principal component 1 (PC1) and y-axis represents principal component 2 (PC2).

volatile compounds detected by GC-MS, and the intensities of 19 sensory descriptors for all fermented and unfermented coffee brews. The PCA of e-nose measurements and volatile compounds (Fig. 3**a**) accounted for 98.40% of the total data variability, with principal component one (PC1) representing 95.57% and PC2 accounting for 2.83%. From the factor loadings in **Table A3**, PC1 is characterized mainly by 2-Furanmethanol, furfural, and acetic acid on the positive side of axis. The PC2 is characterized by 5-methylfurfural, 2-furanmethanol, and maltol on positive

side, methylpyrazine and acetol on negative side. It can be observed that unfermented and fermented green beans (UFTG, UFKG, FTG, and FKG) grouped visibly at the negative side of PC1, which indicated that post-harvest fermentation has no significant impact on both the volatile aromatic compounds of green coffee brew. As for roasted coffee, significant distance could be observed from fermented and unfermented types. Volatile aromatic compounds in the positive quadrant of PC2 could differentiate FTD from UFTD. The hydrolysis of macromolecules and microbial metabolism of carbon and nitrogen occur during wet fermentation could produce essential aroma precursors, including reducing sugar, amino acids, esters, higher alcohols, aldehydes, and ketones, which are related to the coffee aroma after roasting (Bressani et al., 2020; Dzialo et al., 2017; Lee et al., 2015).

Furans, pyrazines, and pyridines are the main groups formed during the thermal processing of the non-volatile precursors present in green coffee beans, including polysaccharides, lipids, protein, and amino acids. 5-Methylfurfural, 2-furanmethanol, and maltol were closer to FTD, which could bring sweet, caramellic, and roasted aroma and flavor.

For Fig. 3**b**, the PCA of e-nose results and sensory descriptors accounted for 87.57% of the total data variability, with the PC1 and PC2 representing 82.54% and 5.03%, respectively. As observation of factor loading in **Table A4**, PC1 was characterized mainly by roasted aroma and flavor, overall quality, and aftertaste on the positive side of the axis, and hay flavor, sweet flavor and aroma on the negative side. On the other hand, the PC2 was characterized by bitter taste on the positive side, and floral aroma on the negative side. UFKG was closer to hay aroma and flavor while UFKL, UFKM, UFKD, and UFTD were relatively closer to sour taste. FKM had more aftertaste. Therefore, spontaneously fermentation could be inferred to influence the generation of non-volatile aroma precursors in coffee beans, which impact could be manifested by intensive roasting in the development of volatile aromatic compounds and sensory perception of fermented and unfermented coffee.

3.3. Machine learning (ML) modelling

NIR and e-nose results from unfermented and fermented coffee were used as inputs in Model 1 and 2, respectively, to predict the peak area of 26 volatile aromatic compounds detected by GC-MS. Table 3 showed that Models 1 and 2 had similar and high overall accuracy ($R = 0.98$; $R = 0.99$) and slope values > 0.90. Both models had no signs of under- or overfitting according to their performance MSE values, where the training-stage MSE values were lower than the values of the validation and testing stages. Moreover, the MSE values of validation and testing stages were close to each other for both models.

Fig. 4 presented the overall regression Models 1 and 2 with 95% prediction bounds. Model 1 had 6.12% outliers (229 out of 3744 data points) according to the 95% prediction bounds, with the majority from ethanol, furfural, and 2-furanmethanol. As for Model 2, it only had 1.53% outliers (191 out of 12,480 data points), where most were from the same aromatic compounds as Model 1. A previous publication from (Gonzalez Viejo et al., 2021) also reported a very high accuracy ($R = 0.99$) in the models developed using e-nose measurements as inputs to predict the peak area of detected volatile compounds by GC-MS in coffee samples with different roasting intensities.

Table 3
The results of the artificial neural network regression Models 1 and 2. Abbreviations: R: correlation coefficient; MSE: means squared error.

Stage	Samples	Observations	R	Slope	Performance (MSE)
Model 1: Inputs: NIR; Targets: volatile aromatic compounds					
Training	100	2,620	0.99	1.00	7.84
Validation	22	562	0.96	0.90	38.94
Testing	22	562	0.97	1.00	34.53
Overall	144	3,744	0.98	1.00	–
Model 2: Inputs: electronic nose; Targets: volatile aromatic compounds					
Training	336	8,736	0.99	1	0.09
Validation	72	1,872	0.99	0.97	14.85
Testing	72	1,872	0.99	0.98	13.17
Overall	480	12,480	0.99	0.99	–

Table 4 showed that Model 3 had high overall accuracy ($R = 0.87$) in predicting 28 sensory descriptors in unfermented and fermented coffee using their NIR absorbance values as inputs without the signs of under- or overfitting. The training MSE value (1.68) was significantly lower than that of the testing (15.42). Model 4, had higher overall accuracy ($R = 0.91$) than Model 3, and also did not have sings of under- or overfitting with the MSE value of the training stage (3.24) lower than the testing stage (22.94).

Fig. 5**a and b** displayed the overall regression Models 3 and 4, respectively with 95% prediction bounds. Model 3 had 4.89% outliers (197 out of 4032 data points), where most outliers were aroma sweet, flavor butter, aftertaste, and astringency mouthfeel. Similarly, there were 3.86% outliers detected in Model 4 (519 out of 13,440 data points), with the majority from butter flavor, followed by taste sourness. Harris et al. (2023) also reported a high accuracy ($R = 0.95$) in an ANN model using NIR absorbance data as inputs to predict the intensity of 20 sensory descriptors from sensory analysis in wine.

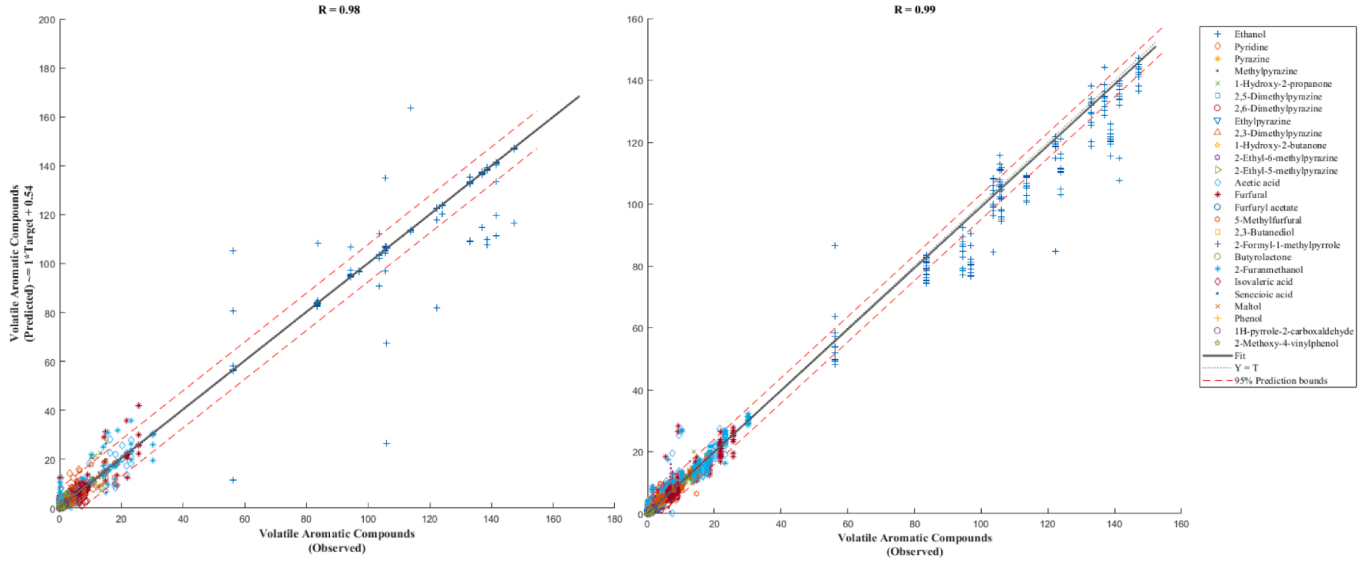
The developed models showed that the NIR (Models 1 and 3) and e-nose (Models 2 and 4) measurements of unfermented and fermented coffee accurately predicted the volatile aromatic compounds and the intensity of the sensory descriptors. Generally, the sensory quality evaluation of coffee is usually conducted through descriptive sensory tests, requiring formal training of sensory panelists, which is time-consuming and costly (Stone et al., 2020). The prediction models constructed using the NIR and e-nose measurements show their higher potential to be applied in specialty coffee processing and its sensory quality assessment for the coffee industry. Furthermore, the low-cost and portable e-nose can also be utilized by local family coffee farms and coffee roasters for process monitoring small coffee retailers for quality assurance as it is more rapid, convenient, portable, affordable, objective, and reliable.

4. Conclusion

Postharvest spontaneous fermentation could only significantly contribute to the liquid color and bitterness of coffee. The differences in perceived intensities for sensory descriptors and the pattern of volatile aromatic compounds of unfermented and fermented coffee brew could be manifested more clearly after intensive roasting. Dark-roasted fermented coffee brew presented higher intensity of perceived liquid color, crema properties, dark chocolate aroma and flavor, roasted aroma and flavor, body, and butter flavor; however, unfermented coffee had a more intense taste of bitterness. The ANN models using the measurements of NIR (Model 1, 3; $R > 0.97$) and e-nose (Models 2, 4; $R > 0.91$) as inputs and GC-MS and QDA test as the ground truth predicted the composition of 26 volatile aromatic compounds and the intensities of 28 sensory descriptors with high accuracy. The combination of novel applications (NIR and e-nose) and machine learning modelling with rapid, reliable, and low-cost properties in the evaluation of coffee physicochemical data and the prediction of sensory properties offers substantial potential in the differentiation and development of specialty coffee and beverages, process monitoring, and quality assurance. Future studies could explore the potential of non-spontaneous fermentation with specific or mixed cultures of yeast and bacterial species for improving coffee quality combined with consumer acceptance and preference.

Funding

This research was funded by the Australian Government as Hafiz Suleria is the recipient of an Australian Research Council Award (ARC-DECRA—DE220100055), the University of Melbourne under the



(a) (b)

Fig. 4. Overall regression Models 1 (a) and 2 (b) predicted 26 volatile aromatic compounds from GC-MS analysis, where the x-axis represents the observed data, while the y-axis depicts the predicted values. Abbreviation R: correlation coefficient.

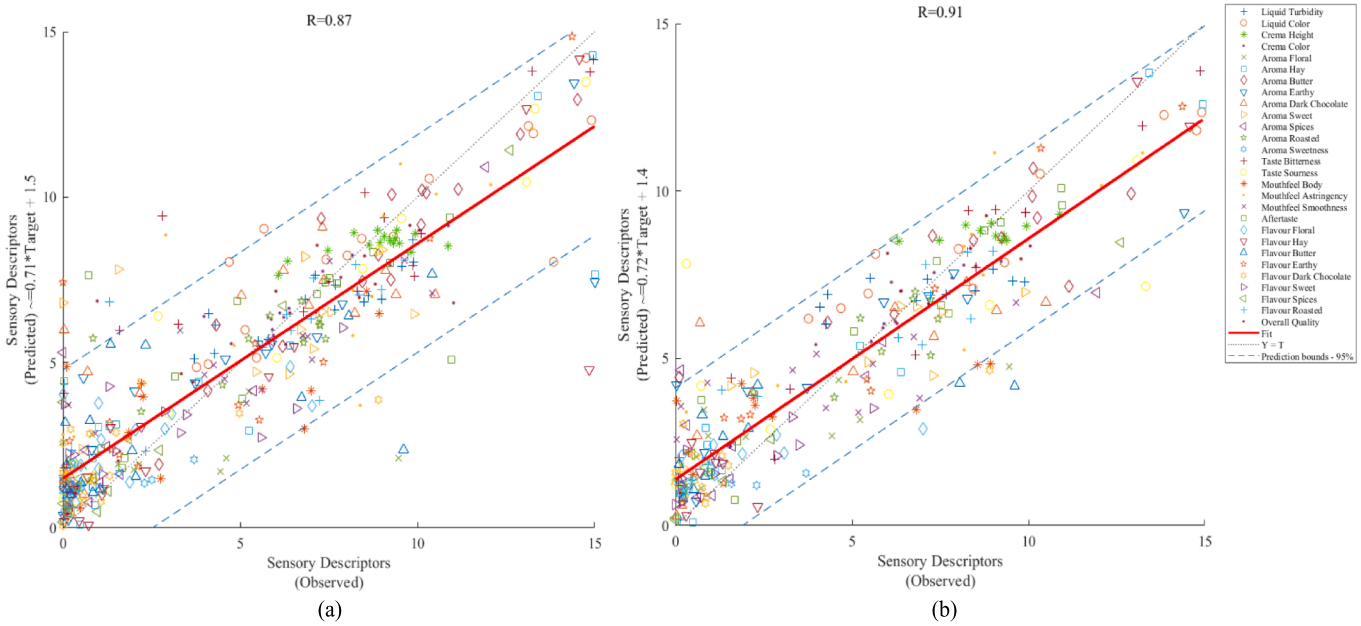
Table 4
The results of the artificial neural network regression Models 3 and 4. Abbreviations: R: correlation coefficient; MSE: means squared error.

Stage	Samples	Observations	R	Slope	Performance (MSE)
Model 3: Inputs: NIR; Targets: sensory descriptors					
Training	100	2,800	0.90	0.82	1.68
Testing	44	1232	0.85	0.74	15.42
Overall	144	4,032	0.87	0.71	–
Model 4: Inputs: electronic nose; Targets: sensory descriptors					
Training	336	9,408	0.93	0.82	3.24
Testing	144	4,032	0.87	0.71	22.94
Overall	480	13,440	0.91	0.72	–

McKenzie Fellowship Scheme (grant no. UoM-18/21), the Future Food Hallmark Research Initiative Funds (grant no. UoM-21/23) and Collaborative. Hanjing Wu is a recipient of a University of Melbourne Research Scholarship provided by The University of Melbourne. Grant reference: 767803.

Ethical Statement

The study protocol was approved by the Human Ethics Advisory Group at the Faculty of Veterinary and Agricultural Sciences, The University of Melbourne, Australia. All experiments were performed in accordance with relevant named guidelines and regulations. All participants provided written informed consent.



(a) (b)

Fig. 5. Overall regression Models 3 (a) and 4 (b) predicted 19 sensory descriptors from quantitative descriptive analysis (QDA), where the x-axis represents the observed data, while the y-axis depicts the predicted values. Abbreviation R: correlation coefficient.

CRediT authorship contribution statement

Hanjing Wu: Conceptualization, Data curation, Formal analysis, Funding acquistion, Investigation, Methodology, Software, Visualization, Writing - original draft preparation, Writing - review & editing. **Claudia Gonzalez Viejo:** Conceptualization, Data curation, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – review & editing. **Sigfredo Fuentes:** Methodology. **Frank R. Dunshea:** Investigation. **Hafiz A.R. Suleria:** Conceptualization, Funding acquisition, Supervision, Visualization, Writing - reviewing & 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.

Data availability

Data will be made available on request.

Acknowledgment

We would like to thank Natalie Harris from the Digital Agriculture, Food and Wine Sciences Group and other Dr. Hafiz Suleria's group researchers from the School of Agriculture, Food and Ecosystem Sciences (SAFES), Faculty of Science, the University of Melbourne, for their incredible support. We also appreciated Australian local coffee farms, Kahawa Estate Coffee and Mountain Top Coffee for green coffee beans for the present research.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodres.2023.113800.

References

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.

Bressani, A. P. P., Batista, N. N., Ferreira, G., Martinez, S. J., Simão, J. B. P., Dias, D. R., & Schwan, R. F. (2021). Characterization of bioactive, chemical, and sensory compounds from fermented coffees with different yeasts species. *Food Research International, 150*, Article 110755.

Bressani, A. P. P., Martinez, S. J., Sarmento, A. B. I., Borém, F. M., & Schwan, R. F. (2020). Organic acids produced during fermentation and sensory perception in specialty coffee using yeast starter culture. *Food Research International, 128*, Article 108773.

Buratti, S., Sinelli, N., Bertone, E., Venturello, A., Casiraghi, E., & Geobaldo, F. (2015). Discrimination between washed Arabica, natural Arabica and Robusta coffees by using near infrared spectroscopy, electronic nose and electronic tongue analysis. *Journal of the Science of Food and Agriculture, 95*(11), 2192–2200.

da Silva Oliveira, E. C., da Luz, J. M. R., de Castro, M. G., Filgueiras, P. R., Guarçoni, R. C., de Castro, E. V. R., da Silva, M. d. C. S., & Pereira, L. L. (2022). Chemical and sensory discrimination of coffee: impacts of the planting altitude and fermentation. *European Food Research and Technology, 248*(3), 659-669.

De Bruyn, F., Zhang, S. J., Pothakos, V., Torres, J., Lambot, C., Moroni, A. V., … De Vuyst, L. (2017). Exploring the impacts of postharvest processing on the microbiota and metabolite profiles during green coffee bean production. *Applied and Environmental Microbiology, 83*(1). e02398-02316.

de Melo Pereira, G. V., de Carvalho Neto, D. P., Júnior, A. I. M., Vásquez, Z. S., Medeiros, A. B., Vandenberghe, L. P., & Soccol, C. R. (2019). Exploring the impacts of postharvest processing on the aroma formation of coffee beans–A review. *Food Chemistry, 272*, 441–452.

Dzialo, M. C., Park, R., Steensels, J., Lievens, B., & Verstrepen, K. J. (2017). Physiology, ecology and industrial applications of aroma formation in yeast. *FEMS Microbiology Reviews, 41*(Supp_1), S95–S128.

Elhalis, H., Cox, J., Frank, D., & Zhao, J. (2021). Microbiological and chemical characteristics of wet coffee fermentation inoculated with hansinaspora uvarum and Pichia kudriavzevii and their impact on coffee sensory quality. *Frontiers in Microbiology, 12*, Article 713969.

Elhalis, H., Cox, J., & Zhao, J. (2022). Coffee fermentation: Expedition from traditional to controlled process and perspectives for industrialization. *Applied Food Research, 100253*.

Fujimoto, H., Narita, Y., Iwai, K., Hanzawa, T., Kobayashi, T., Kakiuchi, M., … Tahara, Y. (2021). Bitterness compounds in coffee brew measured by analytical instruments and taste sensing system. *Food Chemistry, 342*, Article 128228.

Giraudo, A., Grassi, S., Savorani, F., Gavoci, G., Casiraghi, E., & Geobaldo, F. (2019). Determination of the geographical origin of green coffee beans using NIR spectroscopy and multivariate data analysis. *Food Control, 99*, 137–145.

Gonzalez Viejo, C., Fuentes, S., Godbole, A., Widdicombe, B., & Unnithan, R. R. (2020). Development of a low-cost e-nose to assess aroma profiles: An artificial intelligence application to assess beer quality. *Sensors and Actuators B: Chemical, 308*, Article 127688.

Gonzalez Viejo, C., Fuentes, S., Torrico, D., Howell, K., & Dunshea, F. R. (2018a). Assessment of beer quality based on foamability and chemical composition using computer vision algorithms, near infrared spectroscopy and machine learning algorithms. *Journal of the Science of Food and Agriculture, 98*(2), 618–627.

Gonzalez Viejo, C., Fuentes, S., Torrico, D. D., Howell, K., & Dunshea, F. R. (2018b). Assessment of beer quality based on a robotic pourer, computer vision, and machine learning algorithms using commercial beers. *Journal of Food Science, 83*(5), 1381–1388.

Gonzalez Viejo, C., Tongson, E., & Fuentes, S. (2021). Integrating a low-cost electronic nose and machine learning modelling to assess coffee aroma profile and intensity. *Sensors, 21*(6), 2016.

Gonzalez Viejo, C., Torrico, D., Dunshea, F., & Fuentes, S. (2019). Emerging technologies based on artificial intelligence to assess the quality and consumer preference of beverages. *Beverages, 5*(4), 62.

Harris, N., Gonzalez Viejo, C., Barnes, C., & Fuentes, S. (2023). Non-invasive digital technologies to assess wine quality traits and provenance through the bottle. *Fermentation, 9*(1), 10.

Lee, L. W., Cheong, M. W., Curran, P., Yu, B., & Liu, S. Q. (2015). Coffee fermentation and flavor–An intricate and delicate relationship. *Food Chemistry, 185*, 182–191.

Masoud, W., Bjørg Cesar, L., Jespersen, L., & Jakobsen, M. (2004). Yeast involved in fermentation of Coffea arabica in East Africa determined by genotyping and by direct denaturating gradient gel electrophoresis. *Yeast, 21*(7), 549–556.

Nursten, H. E. (2005). *The Maillard reaction: Chemistry, biochemistry and implications*. Royal Society of Chemistry.

Pereira, T. S., Batista, N. N., Pimenta, L. P. S., Martinez, S. J., Ribeiro, L. S., Naves, J. A. O., & Schwan, R. F. (2022). Self-induced anaerobiosis coffee fermentation: Impact on microbial communities, chemical composition and sensory quality of coffee. *Food Microbiology, 103*, Article 103962.

Robinson, R. K. (2014). *Encyclopedia of food microbiology*. Academic press.

Scholz, M. B. d. S., Prudencio, S. H., Kitzberger, C. S. G., & Silva, R. S. d. S. F. d. (2019). Physico-chemical characteristics and sensory attributes of coffee beans submitted to two post-harvest processes. *Journal of Food Measurement and Characterization, 13*, 831-839.

Stone, H., Bleibaum, R., & Thomas, H. A. (2020). *Sensory evaluation practices*. Academic press.

Sualeh, A., Tolessa, K., & Mohammed, A. (2020). Biochemical composition of green and roasted coffee beans and their association with coffee quality from different districts of southwest Ethiopia. *Heliyon, 6*(12), e05812.

Wang, X. (2019). Near-infrared spectroscopy for food quality evaluation. In *Evaluation technologies for food quality* (pp. 105-118). Elsevier.

Wu, H., Lu, P., Liu, Z., Sharifi-Rad, J., & Suleria, H. A. (2022). Impact of roasting on the phenolic and volatile compounds in coffee beans. *Food Science & Nutrition*.

Wu, H., Viejo, C. G., Fuentes, S., Dunshea, F. R., & Suleria, H. A. (2023). The impact of wet fermentation on coffee quality traits and volatile compounds using digital technologies. *Fermentation, 9*(1), 68.

Zamora, M. C., Goldner, M. C., & Galmarini, M. V. (2006). Sourness–sweetness interactions in different media: White wine, ethanol and water. *Journal of Sensory Studies, 21*(6), 601–611.