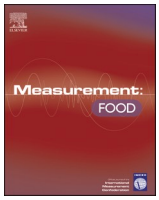
Measurement: Food 14 (2024) 100158

Contents lists available at ScienceDirect

Measurement: Food

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



Predicting dielectric properties of fruit juices at 915 and 2450 MHz using machine learning and physicochemical measurements

Rodrigo Nunes Cavalcanti [a], Vitor Pereira Barbosa [a], Jorge Andrey Wilhelms Gut [a,b], Carmen Cecilia Tadini [a,b,*]

[a] Department of Chemical Engineering, Universidade de São Paulo, Escola Politécnica, São Paulo 05508-010, Brazil
[b] FoRC – Food Research Center, Universidade de São Paulo, São Paulo 05508-080, Brazil

ARTICLE INFO

Keywords:
Microwave heating
Dielectric constant
Loss factor
Electrical conductivity
Dielectric spectroscopy

ABSTRACT

Microwave-assisted thermal processing can provide superior quality for fruit-based products when compared to conventional thermal processing. Understanding the temperature-dependent dielectric properties of liquid foods is needed for the analysis and optimization of the microwave applicator chamber since they govern the heating rate and temperature distribution. While literature offers correlations for specific products, there is a scarcity of methods capable of accommodating variability in composition or predicting behavior for broader product groups. In this study, we measured the dielectric properties (dielectric constant and loss factor) of eight fruit juices (passion fruit, melon, pineapple, cashew, orange, lemon, acerola, and guava) using an open-ended coaxial-line technique for temperatures ranging from 5 to 90 °C at commercial frequencies of 915 and 2450 MHz, alongside electrical conductivity. These properties were successfully correlated with the temperature for each individual juice; then, machine learning techniques (random forest, gradient boosting machine, and multilayer perceptron) were used to predict the properties of this diverse group of eight juices based on various physicochemical measurements. These techniques revealed temperature and electrical conductivity as the most critical predictors, while total solids, pH, acidity, ashes, and select color parameters also emerged as significant variables. These findings demonstrate that the integration of physicochemical analyses with machine learning tools offers an objective approach to correlate and predict dielectric properties for a group of food products, facilitating adjustments in product composition without additional measurements, thus enhancing the efficiency and accuracy of microwave-assisted thermal processing simulations and optimizations.

1. Introduction

Pasteurization is a fundamental processing technique in the food industry in which high temperatures (below 100 °C) are used for the inactivation of microorganisms and enzymes. For processing liquid products such as fruit juices and nectars, heat exchangers are employed for continuous flow heating and cooling of the product stream. One of the drawbacks of thermal processing is quality loss due to the heat [1,2]. Microwave-assisted thermal pasteurization has been thus reported to provide superior quality for fruit-based products when compared to conventional thermal processing [3–8]. The heat exchanger used for heating the product stream to the processing temperature is replaced by a microwave applicator chamber in which the flowing stream absorbs the radiation and heats up [9]. Its advantages are volumetric heating, higher heating rates, better thermal efficiency, and shorter heating times compared to conventional heating methods, resulting in products with better sensorial and nutritional properties [10,11]. The main drawback of microwave heating, mainly in solid and semisolid products, refers to nonuniform temperature distribution; however, this technology has been shown to be suitable for liquid foods, especially in a continuous fluid system since flow can improve thermal mixing [12].

To determine the extent of heating in materials subject to electromagnetic fields, understanding dielectric properties (dielectric constant ε' and dielectric loss factor ε'') is necessary for both designing and optimizing microwave heating systems [13,14]. These properties represent a material's ability to polarize and store electric energy (ε') and the extent of energy dissipation as heat (ε'') [14,15], important for achieving efficient and controlled heating processes. Dielectric properties, influenced by temperature, frequency of the applied electric field, composition, and physicochemical attributes of food [16,17], play a pivotal role in

* Corresponding author at: Department of Chemical Engineering, Universidade de São Paulo, Escola Politécnica, São Paulo 05508-010, Brazil.
E-mail address: catadini@usp.br (C.C. Tadini).

https://doi.org/10.1016/j.meafoo.2024.100158
Received 31 October 2023; Received in revised form 15 March 2024; Accepted 19 March 2024
Available online 20 March 2024

Table 1
Formulation, physicochemical properties and color attributes of the fruit juices.

Names	Fruit juices							
Common name	Yellow passion fruit	'Canary' melon	'Pernambuco' pineapple	Cashew apple	'Pera' sweet orange	'Tahiti' lemon	Acerola cherry	'Paluma' guava
Scientific name	*Passiflora edulis* f. *flavicarpa* Deg.	*Cucumis melo* L.	*Ananas comosus* (L.) Merril	*Anacardium occidentale* L.	*Citrus sinensis* (L.) Osbeck	*Citrus latifolia* Tanaka	*Malpighia emarginata* D.C.	*Psidium guajava* L.
Codification	Passion fruit	Melon	Pineapple	Cashew	Orange	Lemon	Acerola	Guava
Formulation								
Fruit pulp:water ratio	1:6	-	1:2	1:3	-	1:3	1:3	1:3
Sucrose added in juice (g/100 g)	7	-	2	5	-	10	5	3
Analyses								
TS (g/100 g) [1]	8.62 ± 0.03^{e}	11.21 ± 0.01^{b}	13.33 ± 0.02^{a}	8.45 ± 0.02^{f}	9.94 ± 0.07^{d}	10.07 ± 0.09^{c}	7.01 ± 0.02^{h}	7.66 ± 0.04^{g}
TSS (°Brix) [2]	8.63 ± 0.07^{f}	11.66 ± 0.14^{e}	6.68 ± 0.14^{h}	7.83 ± 0.07^{g}	12.83 ± 0.07^{d}	26.23 ± 0.21^{c}	29.62 ± 0.01^{b}	30.27 ± 0.06^{a}
a_w	0.959 ± 0.001^{a}	0.975 ± 0.026^{a}	0.985 ± 0.001^{a}	0.957 ± 0.001^{a}	0.954 ± 0.001^{a}	0.931 ± 0.035^{b}	0.948 ± 0.001^{a}	0.973 ± 0.003^{a}
TTA (g/ 100 mL) [3]	0.538 ± 0.019^{b}	0.081 ± 0.004^{c}	0.163 ± 0.002^{c}	0.057 ± 0.006^{c}	0.637 ± 0.006^{b}	1.356 ± 0.283^{a}	0.600 ± 0.004^{b}	0.111 ± 0.004^{c}
pH	3.06 ± 0.01^{f}	6.28 ± 0.01^{a}	3.97 ± 0.01^{d}	4.86 ± 0.01^{b}	3.96 ± 0.01^{d}	2.44 ± 0.01^{g}	3.47 ± 0.01^{e}	4.09 ± 0.09^{c}
AC (g/100 g) [4] w.b.	0.077 ± 0.002^{d}	0.539 ± 0.030^{b}	0.335 ± 0.001^{c}	0.898 ± 0.107^{a}	0.388 ± 0.001^{c}	0.010 ± 0.039^{d}	0.052 ± 0.004^{d}	0.060 ± 0.001^{d}
Color attributes								
L^*	32.6 ± 0.1^{g}	32.7 ± 0.1^{fg}	35.5 ± 0.1^{e}	65.8 ± 0.1^{a}	42.7 ± 0.5^{c}	33.2 ± 0.1^{f}	39.7 ± 0.1^{d}	52.6 ± 0.1^{b}
a^*	-0.630 ± 0.030^{e}	-0.390 ± 0.030^{d}	-0.780 ± 0.030^{e}	3.88 ± 0.02^{c}	-1.51 ± 0.08^{f}	-0.280 ± 0.040^{d}	22.8 ± 0.1^{a}	18.8 ± 0.1^{b}
b^*	5.80 ± 0.06^{e}	0.440 ± 0.030^{f}	-0.420 ± 0.050^{g}	32.6 ± 0.1^{a}	13.1 ± 0.7^{d}	-1.51 ± 0.07^{h}	16.9 ± 0.1^{b}	15.3 ± 0.1^{c}

All values are expressed as means ± standard deviation of three replicates.
Different letters in the same row mean significant differences at 95 % confidence ($p<0.05$).
[1]*TS* (total solids content). [2]*TSS* (total soluble solids) expressed as °Brix at 20 °C. [3]*TTA* (titratable acidity expressed as citric acid at 25 °C). [4]*AC* (ash content).

microwave heating. In designing and simulating microwave applicator chambers, multi-physics software, such as COMSOL Multiphysics, is indispensable for coupling flow, heat transfer, and electromagnetic propagation phenomena. Such software facilitates the integration of fluid dynamics, heat transfer, and electromagnetic field simulations, allowing for a comprehensive analysis of the complex interactions within the chamber [5,18–22]. The solution of Maxwell's equations of electromagnetism requires knowledge of the dielectric properties' temperature dependence, enabling accurate calculation of heating rates [9,23,24]. Thus, a comprehensive understanding of dielectric properties facilitates not only efficient microwave heating but also enhances control and precision in heating processes.

The scientific literature contains data on the dielectric properties of various foods under specific conditions such as composition, temperature, and frequencies, including different types of milk [16,25–27], fruit juice varieties [22,28,29], coconut water [30], and tomato puree [31], for instance. The published information is mostly limited to correlating measured properties with temperature, or with moisture content for drying applications [16,17,27,32–35]. These correlations are typically established for individual food products, with limited efforts made to generalize them across a broader range of food items [29]. Moreover, only a handful of studies have utilized more advanced prediction techniques, including artificial neural networks [22,25], or multivariate statistical methods [36–38]. Linear regression methods, commonly used for such analyses, may not adequately capture the complexity of the underlying phenomena, as they assume a probabilistic model for data generation. Consequently, algorithmic approaches such as machine learning methods have been proposed as more suitable alternatives [39]. The main challenge arises from the inherent variability in composition and physicochemical attributes among different food products. Each product possesses unique characteristics influenced by factors such as variety, ripeness, growing conditions, and processing methods. This variability complicates the establishment of generalized correlations, making it difficult to accurately capture all nuances since dielectric behavior is intrinsically related to composition. Thus, while single-product correlations provide valuable insights, extending these findings to diverse food groups necessitates careful consideration of the inherent variability across multiple factors.

In this study, we investigate the correlation between the dielectric behavior of eight types of fruit juices and temperature/physicochemical attributes, relevant to microwave-assisted pasteurization processing. By combining data from multiple juices into a single dataset, we aim to determine the interconnectedness of these factors. To achieve this, we employ traditional multiple linear regression (MLR) and advanced machine learning methods such as gradient boosting machine (GBM), random forest (RF), and multilayer perceptron (MLP) for dielectric properties prediction. This study offers a unique contribution by providing a comprehensive analysis of dielectric properties across various juices and comparing the predictive performance of different modeling techniques. Through this analysis, we aim to identify the most influential variables driving dielectric behavior, thereby enhancing our understanding of the mechanisms underlying microwave-assisted pasteurization. MLR, chosen alongside machine learning methods like GBM, RF, and MLP, offers interpretability and simplicity, enabling examination of linear relationships between predictor variables (e.g., temperature, pH, soluble solid content) and dielectric properties. While MLR remains valuable for identifying direct linear associations, machine learning methods are able to capture complex, nonlinear relationships and interactions among predictors, enhancing prediction accuracy in high-dimensional datasets. By integrating both MLR and machine learning methods, this study provides a balanced approach to understanding the relationship between temperature, physicochemical attributes, and dielectric properties in fruit juices.

2. Materials and methods

In summary, several physicochemical attributes were measured for eight fruit beverages: total solids, total soluble solids, water activity, titratable acidity, pH, ashes content and CIELab color. Electrical conductivity and dielectric properties at 915 and 2450 MHz were measured for a range of temperatures compatible with thermal processing. For the data analyses, the dielectric properties were first correlated with the temperature for each beverage individually, and then multilinear regression and machine learning methods were used to correlate these properties with temperature and physicochemical attributes for the group of beverages.

2.1. Raw material and preparation

The eight different types of fruits were purchased from local suppliers in the city of São Paulo - SP (Brazil): yellow passion fruit (*Passiflora edulis* f. *flavicarpa* Deg.), 'Canary' melon (*Cucumis melo* L.), 'Pernambuco' pineapple (*Ananas comosus* (L.) Merril), cashew apple (*Anacardium occidentale* L.), 'Pera' sweet orange (*Citrus sinensis* (L.) Osbeck), 'Tahiti' lemon (*Citrus latifolia* Tanaka), acerola cherry (*Malpighia emarginata*), and 'Paluma' guava (*Psidium guajava* L.). Primarily, the fruits were washed in running water and subsequently sanitized in an aqueous solution of sodium hypochlorite (200 mg L^{-1}) for 30 min. Subsequently, the peels and seeds were manually removed, and the remaining pulps were shredded with a domestic blender RI2101 (Philipps-Wallita, Brazil) and filtered through a stainless-steel household strainer with 1 mm orifices to remove large particles.

The fruit beverages were prepared as ready-to-drink products. Other than melon and orange, the preparation of the remaining six juices involved the addition of Milli-Q® water and food-grade sucrose, with the proportions shown in Table 1. Dilution with water and the addition of sugar aimed to standardize Brix levels and acidity while ensuring optimal palatability. All formulations adhere to U.S. FDA standards (21 CFR 101.30) and international Codex Alimentarius guidelines (Codex Stan 247–2005) for fruit juices and nectars [40,41]. After preparation, the fruit beverages were packed in 300-mL high-density polyethylene bottles and stored at −30 °C in a 349 V Plasma Freezer (FANEM, Brazil).

2.2. Physiochemical characterization

Fruit juice samples were characterized as to total solids (TS), total soluble solids (TSS), titratable acidity (TTA), pH, water activity (a_w) and ashes content (AC) according to official methods AOAC [42]. TSS was determined at room temperature by a 711,849 refractometry (Carl Zeiss Jena, Germany) and corrected according to temperature and acidity [43]. pH and TTA were measured using a pH-Stat PHM-290 (Radiometer, Denmark) and TTA was expressed as the citric acid equivalent. Water activity was measured with AquaLab 3TE (Decagon Devices, USA) at room temperature. Total solids (TS) content was determined in an MA030 vacuum oven (Marconi, Brazil) at 70 ºC and 13.3 kPa. AC was measured as described by method 900.02 in a Q-318D24 muffle furnace (Quimis, Brazil). The color was determined using the CIELab scale, measuring L* (lightness), a* (red-green axis), and b* (yellow-blue axis). RSEX calibration with D65 illuminant and a 10° angle was employed in a ColorQuest XE spectrophotometer (HunterLab, USA).

2.3. Electrical conductivity

The electrical conductivity (EC or σ) of the samples was determined with a YSI3200 conductivity meter and YSI3252 probe (YSI, USA) at the temperature interval (5 to 90) °C. For temperature control, the sample with a thermocouple was immersed in a TC550 thermostatic oil bath (Brookfield, USA).

2.4. Dielectric properties measurement

Usual methods for measuring the dielectric properties of foods at microwave frequencies are open-ended coaxial probe, transmission line, and resonant cavity. The choice depends on the material structure (solid or fluid), frequency range and accuracy. The open-ended coaxial probe is the most common method applied to foods and the most suitable for liquid foods since perfect contact is needed between the probe surface and the sample. This technique requires a network analyzer and is based on the fact that the reflected signal on an open-ended coaxial line attached to a material depends on its dielectric properties [14,44,45].

The reflection coefficient at the probe-sample interface was measured from fruit juice samples with an E5061B vector network analyzer connected to an open-ended coaxial-line probe ("Performance Probe" from the 85070E Dielectric Probe Kit) with a N6314A Type-N test port cable (Agilent Technologies, Malaysia). The measurements were conducted at a temperature interval from (5 to 90) ºC and frequency from (500 to 3000) MHz. The selected temperature range is consistent with juice pasteurization temperatures and the selected frequency range includes the frequencies of industrial interest, which are 915 MHz (wavelength of 0.327 m) and 2450 MHz (wavelength of 0.122 m). These microwave frequencies are among the ISM frequencies, which were allocated by the Federal Communications Commission (FCC) for industrial, scientific and medical electromagnetic energy applications [28].

The dielectric properties (ε' and ε'') were calculated by the 85,070 v. E06.01.36 software (Agilent Technologies, Malaysia) based on the complex reflection coefficient of the sample measured around the probe tip. An electronic calibration module 85093C (Agilent Technologies, Malaysia) was used to minimize interferences and the calibration procedure was performed according to the manufacturer's instructions. The detailed procedure of calibration and dielectric measurements has been reported elsewhere [46].

2.5. Power penetration depth

The penetration depth (d_p) of the microwaves, an important parameter to evaluate the heating uniformity, was calculated according to the following equation [47]:

$$d_p = \frac{c}{2\pi f\sqrt{2\varepsilon'\left[\sqrt{1+\left(\frac{\varepsilon''}{\varepsilon'}\right)^2}-1\right]}} \quad (1)$$

wherein c is the speed of light in free space (2.9979×10^8 m s^{-1}), and f is the electromagnetic field frequency (Hz). The d_p of the microwaves corresponds to the depth at which the power is reduced to $1/e = 36.8$ % (Euler number: $e = 2.7183$) of the incident power at the surface of a semi-infinite body. The complex propagation of the electromagnetic waves in the microwave heating chamber, as affected by the presence of the foodstuff, can be predicted using multi-physics software that solves Maxwell's equations of electromagnetism over the three-dimensional geometry. The penetration depth provides a quick estimate of the heating penetration regarding the foodstuff's main dimensions. Moreover, the dependence of the penetration depth with temperature can suggest if heating would be homogeneous or not along heating time.

2.6. Data analyses

2.6.1. Prediction of the temperature effect on the dielectric properties

Polynomial correlations were adjusted to model the temperature dependence of the dielectric constant, dielectric loss factor, and penetration depth for the commercial frequencies of 915 MHz and 2450 MHz, using TIBCO Statistica 13.4.0 (TIBCO Software, USA). The polynomial order ($n = 1$, 2, or more) was tentatively chosen based on the coefficient

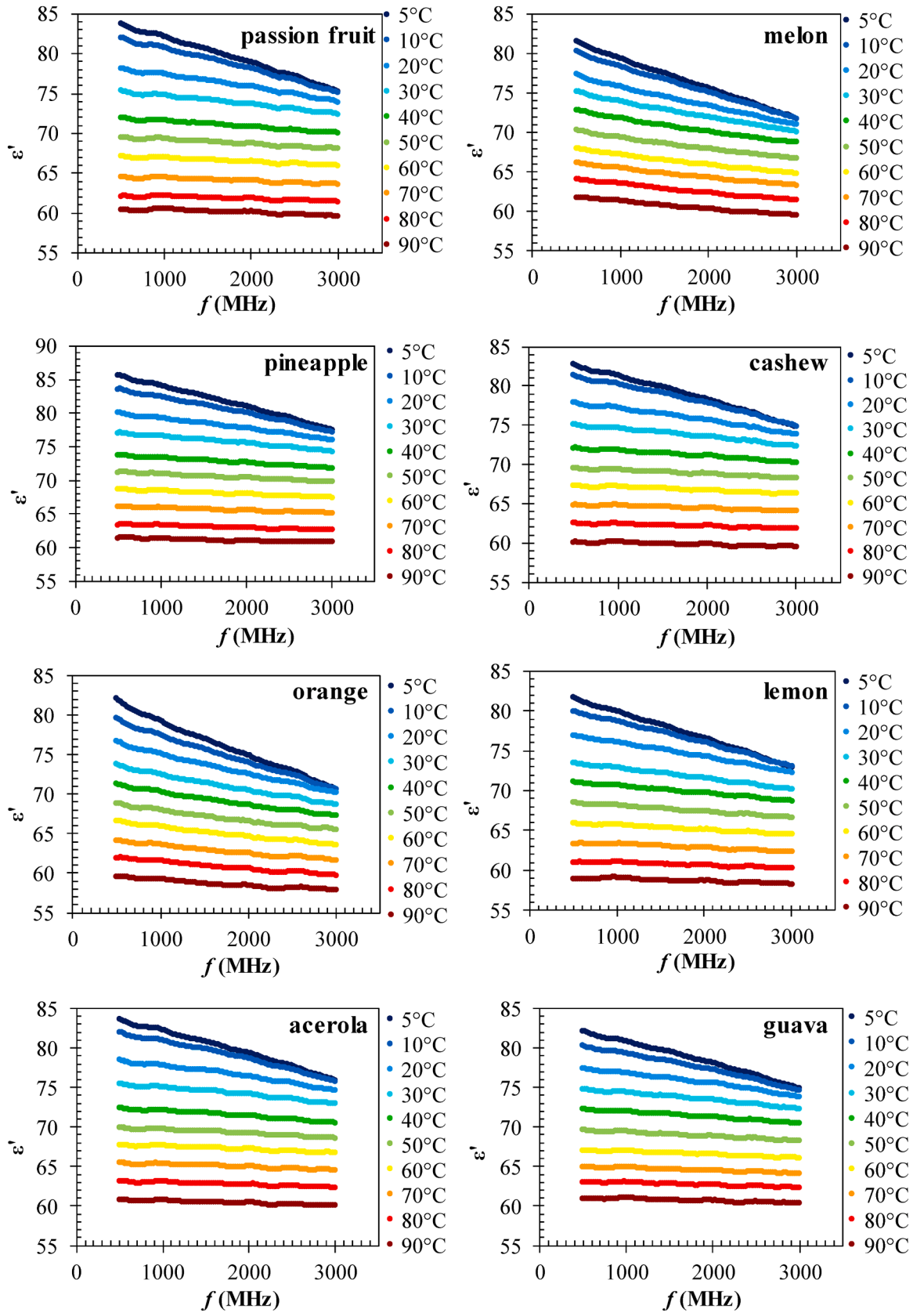


Fig. 1. Dielectric constant (ε′) of the fruit juices measured at a frequency interval from (500 to 3000) MHz and a temperature interval from (5 to 90) °C.

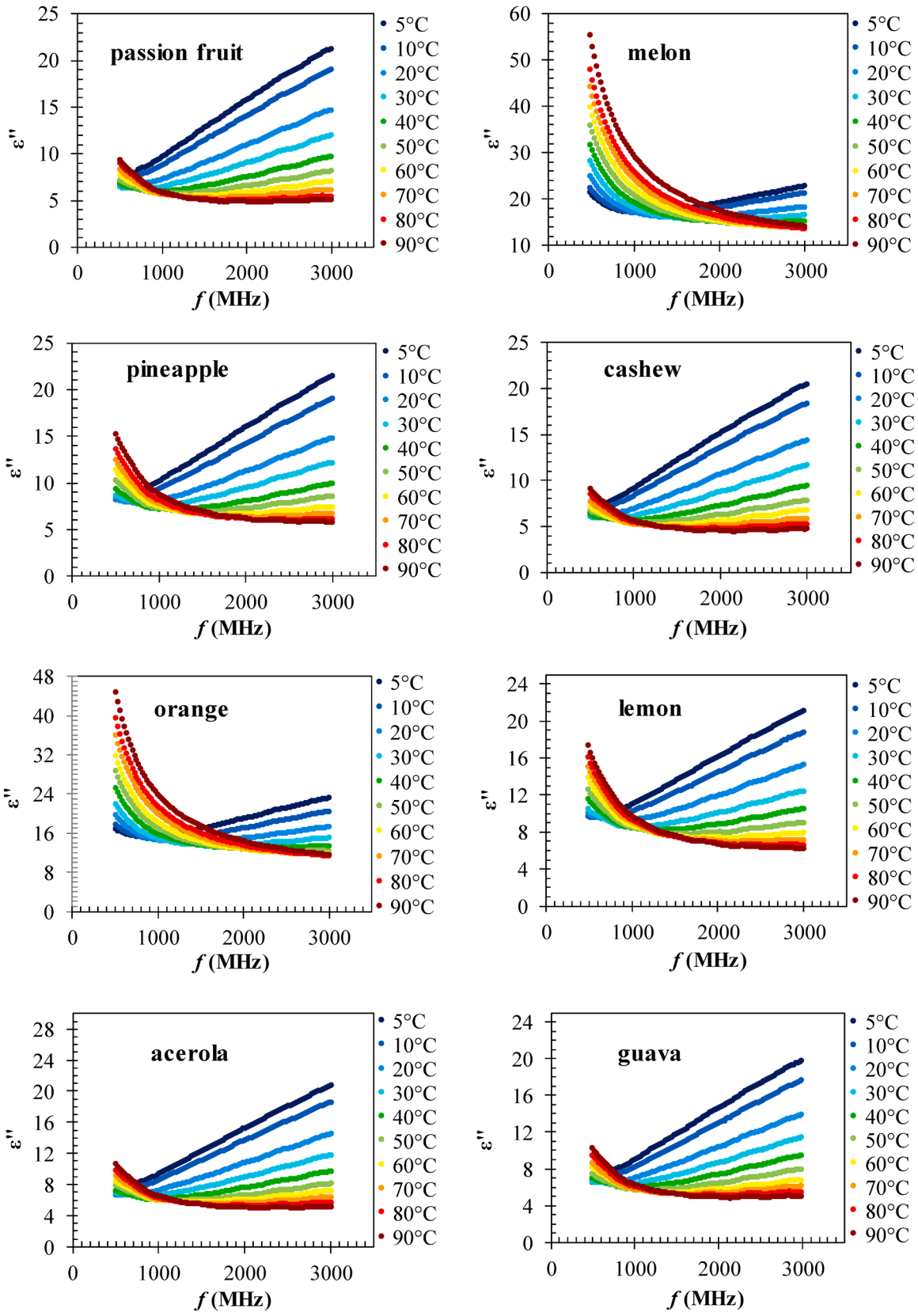


Fig. 2. Dielectric loss factor (ε″) of the fruit juices measured at a frequency interval from (500 to 3000) MHz and a temperature interval from (5 to 90) °C.

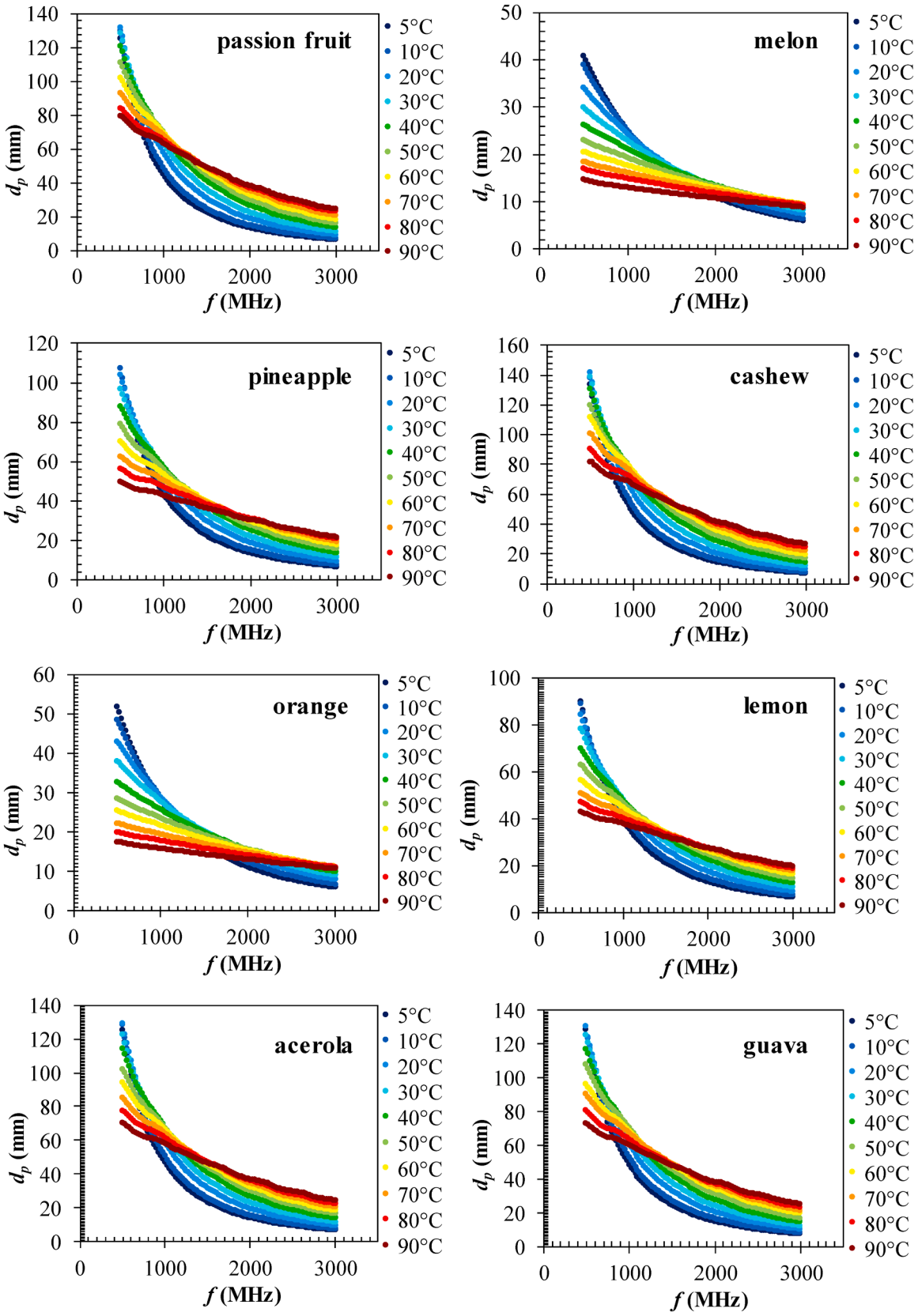


Fig. 3. Penetration depth (d_p) of the fruit juices determined at a frequency interval from (500 to 3000) MHz and a temperature interval from (5 to 90) °C.

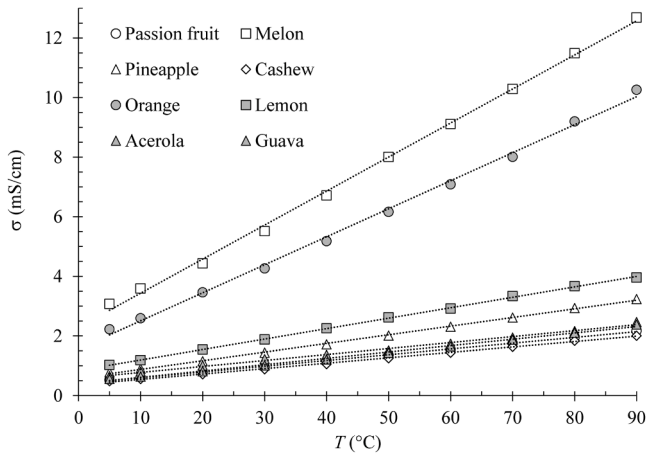


Fig. 4. Experimental data and respective linear regression of the electrical conductivity (σ) of the juices prepared from eight different types of fruit measured at a temperature interval from (5 to 90) °C.

Table 2
Linear regression parameters, standard errors of estimate (δ_{est}) and the coefficient of determination (R^2) of the electrical conductivity ($\sigma = a_0 + a_1 T$) of the fruit juices, valid for the temperature interval from (5 to 90 °C).

Fruit juices	a_0 (mS cm^{-1})	a_1 (mS cm^{-1} °C^{-1})	δ_{est} (-)	R^2
Passion fruit	0.404 ± 0.005	0.0193 ± 0.0001	0.02	0.9992
Melon	2.291 ± 0.047	0.1142 ± 0.0009	0.14	0.9983
Pineapple	0.588 ± 0.010	0.0290 ± 0.0002	0.03	0.9987
Cashew	0.360 ± 0.007	0.0181 ± 0.0001	0.02	0.9985
Orange	1.563 ± 0.051	0.0941 ± 0.0010	0.15	0.9971
Lemon	0.843 ± 0.008	0.0350 ± 0.0002	0.02	0.9995
Acerola	0.577 ± 0.039	0.0199 ± 0.0007	0.11	0.9641
Guava	0.398 ± 0.013	0.0213 ± 0.0002	0.04	0.9983

All values are expressed as means ± standard deviation (n = 3).

of determination (R^2) of each equation [34,46].

2.6.2. Prediction of the multivariable effect on the dielectric properties

In the Multiple linear regression (MLR) analysis, interpretive variables (X_i) such as *T, TSS, a_w, TTA*, pH, *TS, AC, L**, *a**, *b**, and σ were examined for their tentatively linear relationship with response variables (Y_j), which were ε', ε'', and d_p. Model performance was evaluated using R^2 to assess explained variance, RMSE for predictive accuracy, and the significance of coefficients ($p < 0.05$) to identify impactful variables.

Nonlinear correlations between dielectric and physicochemical properties were assessed by using the machine learning methods random forest (RF), gradient boosting machine (GBM), and multilayer perceptron (MLP). The variable importance was determined for each method on a 0–100 scale. Two different datasets were evaluated separately for 915 MHz and 2450 MHz, each one composed of a matrix with 240 rows (samples) and 14 columns (11 independent and 3 response variables). For all the machine learning methods, both datasets were randomly divided into a training set (70 %) and a test set (30 %). The training set was used for fitting the model parameters; then, the adjusted model was used to predict the responses for the data in the test set to evaluate the predictive ability of the model. The RMSE of the prediction dataset was used as a criterion to evaluate the model performance. All the methods (MLR, RF, GBM, MLP) were performed by TIBCO Statistica 13.4.0 (TIBCO Software, USA).

RF can be summarized as an ensemble of decision trees (DT) created during the training step and outputting the mean prediction of the individual trees [48]. Each tree grows based on a bootstrap sampling from the original data, which relies on random sampling with replacement [49], and a subset of the explanatory variables is randomly selected at each node. The number of variables available for splitting at each tree node, the number of trees in the forest, and the node size are the main parameters that affect the stability and sensitivity of the model [50].

GBM is also a self-learning DT algorithm that improves the performance of the regression tree by adopting the gradient boosting algorithm [51,52]. During the training step, an initial regression tree is created, and then the next regression tree is trained taking into account the residual of the previous regression tree; finally, using multiple iterations, a model with high accuracy for predictive results is obtained [52,53].

MLP is a feed-forward neural network comprising an input layer to pass the input vector to the network, one or more hidden layers to perform computations, and an output layer for outputting the responses. After multiple iterations in the training step, the neural network model determines the mathematical functions and weights that correlate input and output data and creates an internal model that can be used to predict new input data. The model is calculated by the interconnection of neurons, and the accuracy of the model is affected by the architecture of the neural network [52].

2.7. Statistical analyses

Experimental data were evaluated by the analysis of variance (ANOVA) followed by Tukey's post hoc test at the 95 % significance level using the software TIBCO Statistica 13.4.0 (TIBCO Software, USA).

3. Results and discussion

3.1. Physicochemical characteristics

Table 1 exhibits the formulation, physicochemical properties, and some compositional aspects of interest of the eight fruit juices. As expected, there were significant differences for total solids (7.01 - 13.3) g/100 g of juice, total soluble solids (6.68 - 30.3) °Brix, water activity (0.931 −0.985), titratable acidity (0.057 - 1.356) g/ 100 g expressed as citric acid, pH (2.44 - 6.28), ashes content (0.010 - 0.898) g/100 g, and color parameters *L** (32.6 - 65.8), *a** (−1.51 - 22.8) and *b** (−1.51 - 32.6). The high variability of those properties is not only because different fruits were used but also due to different formulations applied to each juice. For instance, melon and orange juices were prepared with no addition of water and sucrose, while passion fruit juice was diluted with water at a proportion of 1:6 (v/v) and added with 7 g/100 g of sucrose. This heterogeneity of data enables higher power to generalize the models obtained from machine learning techniques.

3.2. Dielectric properties

Diagrams of the dielectric constant (ε'), dielectric loss factor (ε''), and penetration depth (d_p) of all juices are shown in Figs. 1, 2 and 3, respectively. The profiles of ε' were quite similar among the eight juices, indicating that water is the main factor for the electrical polarization (Fig. 1). According to Table 1, the water content varies between (86.7 and 93.0) g/100 g. The decrease of ε' values with an increase in temperature and frequency can be explained by the reduced polarization under thermal agitation and the reduced dipole response under higher frequencies, respectively [29]. Although ε' values were close for all solutions, distilled water presented a higher dielectric constant [30]. The loss factor ($\varepsilon'' = \varepsilon''_\sigma + \varepsilon''_d$) can be described equally by the influence of two main mechanisms, dipole loss (ε''_d) and ionic loss ($\varepsilon''_\sigma = \sigma / 2\pi f \varepsilon_0$). Fig. 2 allows observation that a rise in temperature may cause an increase or decrease in the loss factor, depending on which the dominant mechanism is [54]. Ionic loss is favored by higher temperatures, which cause an enhancement of ionic motion, and lower frequencies, which are associated with an increase in the length of ionic motion. Overall, this behavior can be better observed at frequencies lower than 1000 MHz. In contrast, the dipole loss is related to the dipolar relaxation of the water

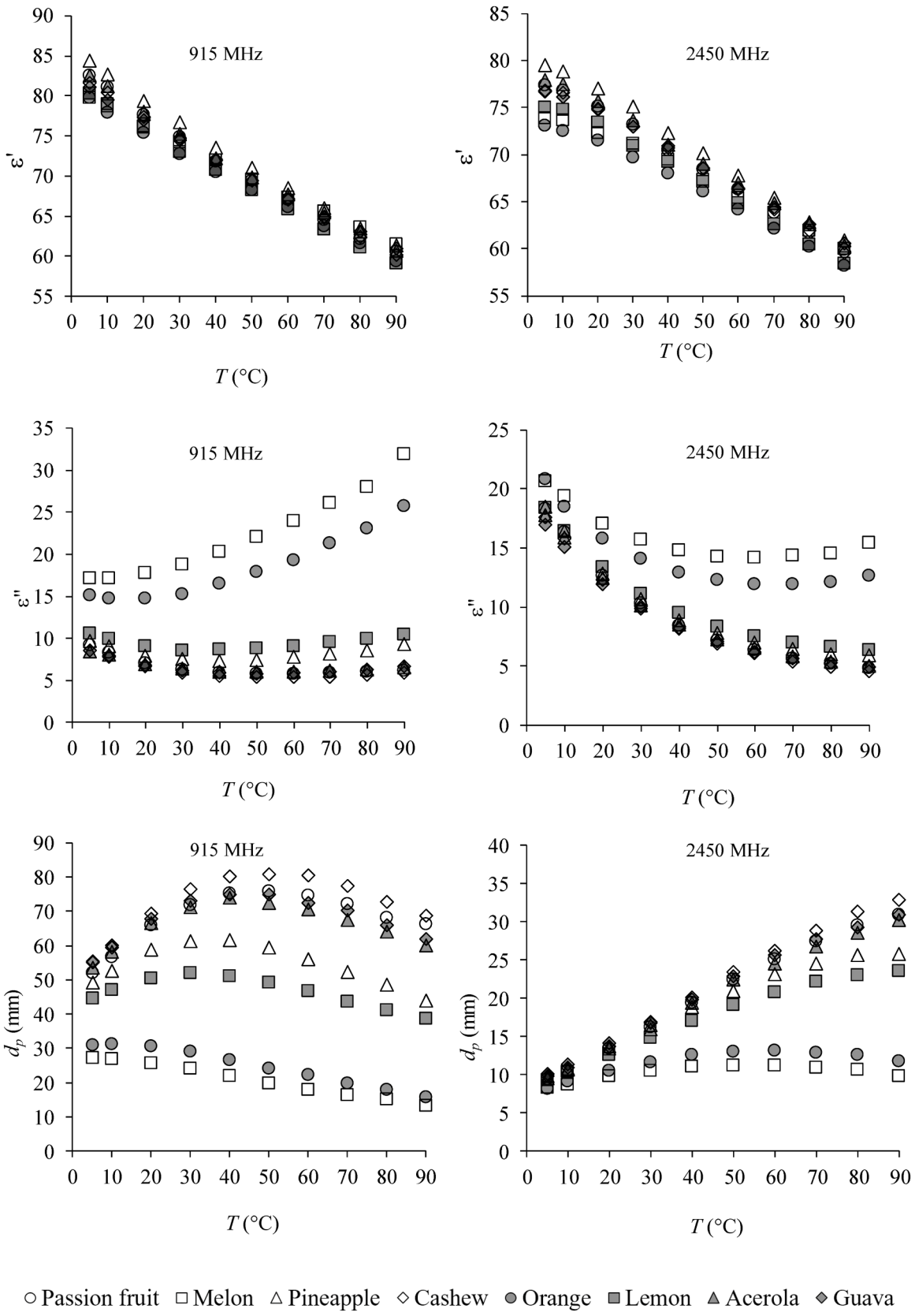


Fig. 5. Dielectric constant (ε′), dielectric loss factor (ε″) and penetration depth (d_p) at 915 MHz and 2450 MHz of the juices prepared from eight different types of fruit measured at a temperature interval from (5 to 90) °C.

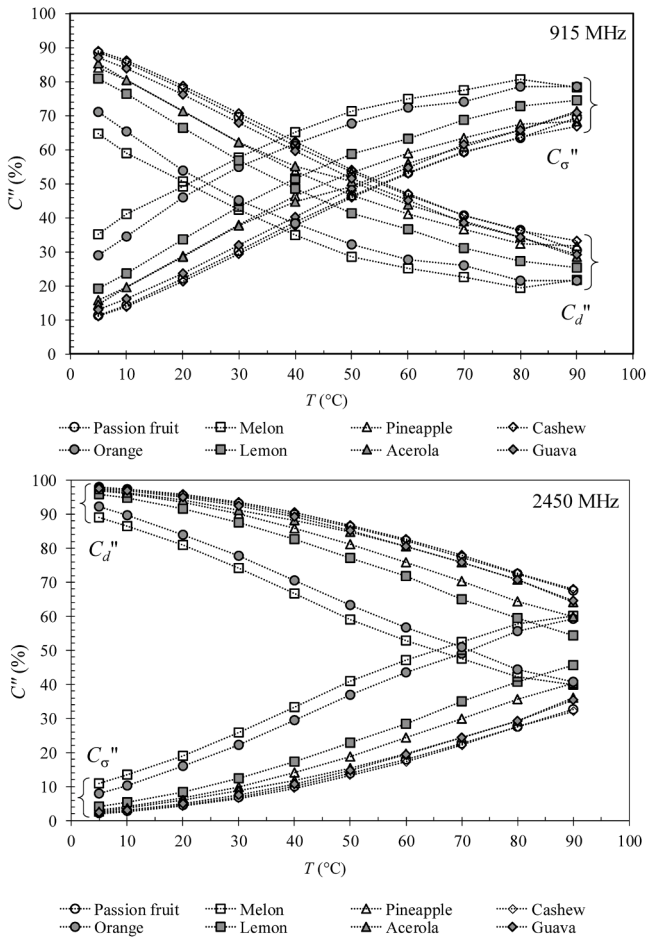


Fig. 6. Contribution of the ionic conduction ($C_\sigma = (\varepsilon''_\sigma/\varepsilon'') \times 100\%$, ○) and dipole rotation ($C_d = (\varepsilon''_d/\varepsilon'') \times 100\%$,) mechanisms on dielectric loss factor (ε'') at 915 and 2450 MHz of the juices prepared from eight different types of fruit measured at a temperature interval from (5 to 90) °C.

molecule, in which the loss factor increases with frequency can be seen for the lower temperatures ($T < 30$ °C). In summary, the behavior observed in the loss factor plot can be considered the transition from an ionic-governed loss to a dipolar-governed loss [30].

On the whole, the penetration depth of the microwaves Eq. (1) decreases with increasing frequency; however, there is a clear change in temperature dependence with increasing frequency (Fig. 3). At lower frequencies, the penetration depth decreases with increasing temperature, while at higher frequencies, the penetration depth increases with temperature. The second behavior is more desirable because, as the food is heated, the radiation penetrates more deeply, promoting more homogenous heating. In the case of low frequencies, the penetration depth decreases with heating, which can cause cold spots in the material. In any case, note that the greater depths are found at the lower frequencies and that the heating pattern of the material depends heavily on its dimensions and geometry.

3.3. Electrical conductivity

As shown in Fig. 4, the higher the temperature, the higher the electrical conductivity vis-à-vis the continuous intensification in the mobility of the ions due to the reduction in the solution viscosity as a result of the progressive increase in temperature [46]. The electrical conductivity exhibited a good linear correlation with temperature ($0.9641 \leq R^2 \leq 0.9995$), as shown in Table 2. Although cashew juice showed the highest content of ashes and, consequently, more salts (Table 1), melon juice was the one with the highest angular coefficient (a_1) and linear coefficient (a_0). This phenomenon indicates that even though the ash content has a positive influence on the electrical conductivity, the addition of sugar to the formulation of cashew juice has promoted a great deleterious effect on the electrical conductivity. Hence, there is a great likelihood that ash content is not the main factor influencing electrical conductivity. Generally, these results agree with the behavior of other liquid foods regarding temperature dependence [16,26,30,46,29,55].

3.4. Commercial frequencies of 915 and 2450 MHz

Fig. 5 shows ε' and ε'' of the juices as a function of temperature at 915 and 2450 MHz commercial frequencies. All eight fruit juices presented a similar behavior with ε' decreasing almost linearly with the temperature at both frequencies. Lower values were obtained for orange, lemon, and melon juices while higher ε' values were achieved by pineapple, acerola, and passion fruit juices. Also observed is that the dielectric loss factor (ε'') showed a nearly constant value at 915 MHz for all juices, except for melon and orange juices, which exhibited an increase of ε'' with temperature. Likewise, all the juices presented a decrease of ε'' with the temperature at 2450 MHz but melon and orange juices suffered a fairly less pronounced decrease. In addition, melon and orange juices had the highest ε'' values at both frequencies. Fig. 5 also exhibits the penetration depth of the eight fruit juices as a function of temperature. At 915 MHz, the d_p values of all samples were higher at intermediate temperatures (40–70 °C) whereas a slight decrease is observed for melon and orange juices as the temperature rises. At 2450 MHz, though, the juices exhibited an increase of d_p with temperature, except melon and orange juices, which had practically constant values of d_p throughout the temperature range studied. Once again, note that melon and orange juices displayed quite distinct behavior. The main reason is undoubtedly the fact that were formulated without adding sugar or water, but no concrete explanation of how this interference occurs can be inferred from the results presented so far. Despite that, it is deducible that more heat can be dissipated from melon and orange juices due to the higher values of loss factor; more uneven heating is attained because of their lower penetration depth.

The dielectric loss factor is a result of the contribution of ionic conduction and dipole rotation mechanisms. Fig. 6 exhibits the relative contributions of ionic conduction, at 915 MHz and 2450 MHz: the ionic conduction ($C_\sigma = (\varepsilon''_\sigma/\varepsilon'') \times 100$) and dipole rotation ($C_d = (\varepsilon''_d/\varepsilon'') \times 100$). As observed, the ionic conduction mechanism is predominant as temperature increases, for both frequencies. This phenomenon was more pronounced for all juices at 915 MHz and for melon and orange juices at 2450 MHz, which might be associated with the higher electrical conductivity values of those juices (Fig. 1). A similar behavior has been reported in the literature [30,46,54].

3.5. Data correlation

3.5.1. Temperature-dependence effect on the dielectric properties

Polynomial correlations adjusted to the temperature dependence for electrical conductivity (σ) are presented in Table 2, and those for dielectric properties (ε', ε'' and d_p) are presented in Table 3. Model fitting was made using the least number of polynomial coefficients to provide a good fit, valid for the range of temperature from (5 to 90) °C. In Table 3, for dielectric constant (ε'), for the regression with linear and quadratic coefficients, the standard errors of estimate (δ_{est}) were lower than 0.54, and a good fit was accomplished for all juices ($0.9920 \leq R^2 \leq 0.9993$). For the dielectric loss factor (ε''), good correlations were obtained for cubic regressions ($0.9824 \leq R^2 \leq 0.9994$), with standard errors of estimate (δ_{est}) lower than 0.48. For penetration depth (d_p), the best adjustment was obtained for quadratic equations ($0.9120 \leq R^2 \leq 0.9984$; $0.01 \leq \delta_{est} \leq 3.3$).

Table 3
Polynomial regression parameters, standard errors of estimate (δ_{est}) and the coefficient of determination (R^2) of the dielectric constant (ε'), dielectric loss factor (ε'') and the penetration depth (d_p) of the fruit juices at 915 MHz and 2450 MHz, valid for the temperature interval from (5 to 90°C).

Juices		f (MHz)	a_0 (× 10)	a_1 (× 10^{-1}°C^{-1})	a_2 (× 10^{-4}°C^{-2})	a_3 (× 10^{-5}°C^{-3})	δ_{est} (-)	R^2
Passion fruit	ε'	915	8.42 ± 0.02	-3.38 ± 0.09	8.30 ± 0.90	-	0.26	0.9989
		2450	7.88 ± 0.02	-1.96 ± 0.11	-2.04 ± 1.15	-	0.34	0.9970
	ε''	915	1.00 ± 0.01	-2.04 ± 0.05	30.5 ± 1.2	-1.41 ± 0.08	0.05	0.9982
		2450	2.07 ± 0.01	-4.85 ± 0.11	54.9 ± 2.8	-2.28 ± 0.20	0.13	0.9994
	d_p (mm)	915	4.83 ± 0.080	10.3 ± 0.412	-94.5 ± 4.29	-	2.40	0.9627
		2450	0.725 ± 0.019	3.38 ± 0.097	-7.81 ± 1.01	-	0.13	0.9978
Melon	ε'	915	8.11 ± 0.03	-2.53 ± 0.18	4.20 ± 1.87	-	0.55	0.9920
		2450	7.50 ± 0.03	-1.37 ± 0.17	-3.62 ± 1.87	-	0.29	0.9883
	ε''	915	1.69 ± 0.04	0.200 ± 0.43	14.8 ± 10.6	0.123 ± 0.741	0.48	0.9919
		2450	2.22 ± 0.02	-3.16 ± 0.17	37.4 ± 4.4	-1.27 ± 0.30	0.20	0.9942
	d_p (mm)	915	2.02 ± 0.012	3.60 ± 0.060	-32.6 ± 0.624	-	0.05	0.9937
		2450	0.753 ± 0.004	1.35 ± 0.022	-12.2 ± 0.233	-	0.01	0.9937
Pineapple	ε'	915	8.59 ± 0.01	-3.30 ± 0.07	6.35 ± 0.72	-	0.21	0.9993
		2450	8.09 ± 0.02	-2.02 ± 0.09	-2.71 ± 0.98	-	0.29	0.9980
	ε''	915	1.05 ± 0.01	-1.76 ± 0.07	28.5 ± 1.8	-1.15 ± 0.13	0.08	0.9917
		2450	2.08 ± 0.01	-4.79 ± 0.12	55.3 ± 3.1	-2.30 ± 0.21	0.14	0.9992
	d_p (mm)	915	4.80 ± 0.081	6.17 ± 0.417	-75.5 ± 4.35	-	2.50	0.9320
		2450	0.727 ± 0.019	3.50 ± 0.096	-15.5 ± 1.00	-	0.13	0.9966
Cashew	ε'	915	8.31 ± 0.02	-2.90 ± 0.11	4.01 ± 1.12	-	0.33	0.9979
		2450	7.79 ± 0.02	-1.57 ± 0.11	-5.45 ± 1.13	-	0.33	0.9969
	ε''	915	0.942 ± 0.006	-1.85 ± 0.06	26.0 ± 1.6	-1.09 ± 0.11	0.07	0.9965
		2450	1.98 ± 0.01	-4.53 ± 0.13	49.8 ± 3.3	-2.04 ± 0.23	0.15	0.9991
	d_p (mm)	915	2.02 ± 0.069	9.14 ± 0.353	-16.0 ± 3.69	-	1.80	0.9965
		2450	0.753 ± 0.026	3.41 ± 0.132	-5.98 ± 1.38	-	0.25	0.9965
Orange	ε'	915	8.08 ± 0.02	-2.70 ± 0.12	3.81 ± 1.26	-	0.37	0.9969
		2450	7.36 ± 0.03	-0.78 ± 0.31	-18.98 ± 7.79	-	0.35	0.9961
	ε''	915	1.53 ± 0.03	-0.781 ± 0.259	29.9 ± 6.4	-0.950 ± 0.449	0.29	0.9949
		2450	2.26 ± 0.02	-4.34 ± 0.17	56.5 ± 4.2	-2.38 ± 0.29	0.19	0.9969
	d_p (mm)	915	3.25 ± 0.035	-1.09 ± 0.179	-8.93 ± 1.86	-	0.45	0.9862
		2450	0.720 ± 0.004	1.95 ± 0.019	-16.1 ± 0.195	-	0.01	0.9984
Lemon	ε'	915	8.18 ± 0.02	-2.97 ± 0.10	4.94 ± 1.07	-	0.31	0.9981
		2450	7.63 ± 0.02	-1.65 ± 0.13	-4.08 ± 1.31	-	0.39	0.9956
	ε''	915	1.12 ± 0.01	-1.61 ± 0.09	28.9 ± 2.3	-1.34 ± 0.16	0.10	0.9824
		2450	2.04 ± 0.02	-4.31 ± 0.16	47.7 ± 4.1	-1.93 ± 0.28	0.19	0.9984
	d_p (mm)	915	4.44 ± 0.069	3.40 ± 0.354	-46.8 ± 3.69	-	1.80	0.9120
		2450	0.749 ± 0.013	2.91 ± 0.067	-12.4 ± 0.694	-	0.06	0.9977
Acerola	ε'	915	8.41 ± 0.02	-3.18 ± 0.10	6.76 ± 1.08	-	0.32	0.9983
		2450	7.93 ± 0.02	-1.89 ± 0.11	-2.63 ± 1.13	-	0.33	0.9973
	ε''	915	0.965 ± 0.066	-1.85 ± 0.06	28.6 ± 1.6	-1.30 ± 0.11	0.07	0.9948
		2450	2.00 ± 0.01	-4.62 ± 0.14	53.1 ± 3.5	-2.26 ± 0.24	0.16	0.9989
	d_p (mm)	915	5.08 ± 0.093	9.01 ± 0.480	-90.9 ± 5.00	-	3.30	0.9288
		2450	0.770 ± 0.019	3.30 ± 0.098	-8.72 ± 1.02	-	0.14	0.9975
Guava	ε'	915	8.25 ± 0.03	-2.85 ± 0.15	5.13 ± 1.57	-	0.46	0.9959
		2450	7.80 ± 0.03	-1.70 ± 0.18	-3.02 ± 1.83	-	0.54	0.9920
	ε''	915	0.916 ± 0.079	-1.62 ± 0.08	23.9 ± 1.9	-1.01 ± 0.13	0.09	0.9914
		2450	1.9 ± 0.02	-4.22 ± 0.16	46.7 ± 3.9	-1.90 ± 0.27	0.18	0.9984
	d_p (mm)	915	5.25 ± 0.088	9.10 ± 0.454	-91.4 ± 4.73	-	2.90	0.9371
		2450	0.801 ± 0.024	3.39 ± 0.123	-9.04 ± 1.28	-	0.21	0.9962

All values are expressed as means ± standard deviation.

3.5.2. Multivariable effect on the dielectric properties

The MLR, GBM, RF, and MLP methods were applied to predict dielectric constant (ε'), dielectric loss factor (ε'') and penetration depth (d_p), each one as a function of temperature, soluble solids (*TSS*), water activity (a_w), titratable acidity (*TTA*), pH, total solids (*TS*), ashes content (*AC*), electrical conductivity (σ), and color parameters L^*, a^* and b^*. Prediction performances were evaluated using R^2 and/or RMSE between predicted and experimental values as criteria.

In Table 4, it can be noticed that MLR achieved a good fit with observed data for the majority of the dielectric properties at both frequencies ($R^2 \geq 0.982$ and RMSE ≤ 0.936), except for ε'' at 915 MHz ($R^2 = 0.790$; RMSE = 2.16) and 2450 MHz ($R^2 = 0.895$; RMSE = 1.47), and d_p at 915 MHz ($R^2 = 0.782$; RMSE = 8.51). Temperature and electrical conductivity were considered significant parameters ($p < 0.05$) for all the response variables at both electric field frequencies. All the parameters were considered significant for the dielectric constant at both frequencies. For the dielectric loss factor at 915 MHz, only pH and ashes content were considered non-significant while only temperature and electrical conductivity were significant at 2450 MHz. For penetration depth, titratable acidity, pH, and ashes content were considered non-significant at 915 MHz whereas only the constant was omitted in the equation at 2450 MHz. For each response variable, better fits were found at 2450 MHz, indicating this frequency might be a better choice than 915 MHz for predicting dielectric properties by MLR. Similar trends were found for the MLR of sheep milk [56].

The GBM model needs to tune several hyper-parameters to optimize the model performance. The bag fraction of 50 % stipulates the percentage of the training set applied to each interaction, while the maximum number of trees (2000), and the minimum number of cases in a node (1) were used as a stop criterion for the GBM model. Different learning rates were used (0.01, 0.05 and 0.1), in which diverse values of minimum node size (1–5), the maximum number of nodes in each tree (3–15), and the maximum number of levels in a tree (depth of the tree) (3–30) were assessed during training. The latter aforementioned parameters are responsible for the tree complexity, limiting the maximum number of interactions between predictors. GBM models performed better with a minimum node size of 1, and a maximum of 30 levels for each tree, as well as a learning rate not lower than 0.1. The optimal

Table 4
Parameters of the multiple linear regression (MLR), expressed as $Y_i = \beta_0 + \sum_{i=1}^{n} \beta_i X_i$, standard errors of estimate (δ_{est}), coefficients of determination (R^2) and root mean square error (RMSE) for the temperature, physicochemical attributes and composition dependence of the dielectric properties: dielectric constant (ε'), dielectric loss factor (ε'') and penetration depth (d_p).

X_i	β_i	ε'		ε"		d_p (mm)	
		915 MHz	2450 MHz	915 MHz	2450 MHz	915 MHz	2450 MHz
Constant	β_0	86.6 ± 4.62	99.8 ± 3.49	-49.1 ± 11.8	9.00 ± 8.00*	205 ± 46.4	-8.25 ± 5.11*
T	β_1	-1.06 ± 0.013	-1.03 ± 0.011	-0.363 ± 0.045	-1.01 ± 0.032	0.359 ± 0.046	1.17 ± 0.013
TSS [1]	β_2	-0.579 ± 0.086	-1.04 ± 0.077	1.95 ± 0.311	0.212 ± 0.220*	-1.47 ± 0.317	0.484 ± 0.090
a_w	β_3	0.003 ± 0.015	-0.075 ± 0.014	0.245 ± 0.055	0.039 ± 0.039*	-0.120 ± 0.056	0.048 ± 0.016
TTA [2]	β_4	0.073 ± 0.050	0.325 ± 0.044	-0.745 ± 0.181	-0.092 ± 0.127*	0.353 ± 0.184*	-0.169 ± 0.052
pH	β_5	-0.373 ± 0.034	-0.158 ± 0.030	-0.077 ± 0.122*	0.001 ± 0.086*	-0.071 ± 0.124*	0.260 ± 0.035
TS [3]	β_6	-0.040 ± 0.030	-0.204 ± 0.026	0.773 ± 0.107	0.039 ± 0.076*	-0.758 ± 0.109	0.094 ± 0.031
AC [4]	β_7	0.093 ± 0.032	0.175 ± 0.029	-0.164 ± 0.116*	0.024 ± 0.082*	0.062 ± 0.118*	-0.096 ± 0.034
*L**	β_8	0.419 ± 0.080	0.830 ± 0.070	-1.78 ± 0.287	-0.258 ± 0.202*	1.22 ± 0.292	-0.491 ± 0.083
*a**	β_9	0.693 ± 0.080	1.14 ± 0.071	-1.99 ± 0.287	-0.222 ± 0.203*	1.43 ± 0.292	-0.512 ± 0.083
*b**	β_{10}	-0.603 ± 0.101	-1.24 ± 0.090	2.82 ± 0.365	0.279 ± 0.257*	-1.97 ± 0.371	0.623 ± 0.106
σ [5]	β_{11}	0.211 ± 0.021	0.138 ± 0.019	0.568 ± 0.077	0.561 ± 0.054	-0.506 ± 0.078	-0.980 ± 0.022
δ_{est}		0.869	0.656	2.22	1.50	8.73	0.960
R^2		0.984	0.987	0.790	0.895	0.782	0.982
RMSE		0.848	0.639	2.16	1.47	8.51	0.936

All values are expressed as means ± standard deviation.
* Non-significant parameter at the 95 % significance level ($p>0.05$). [1]*TSS* (total soluble solids) are expressed as °Brix at 20°C; [2]*TTA* (titratable acidity) expressed as citric acid at 25°C; [3]*TS* (total solids content); [4]*AC* (ashes content); [5]σ (electrical conductivity).

Table 5
Performance of the machine learning methods GBM (gradient boosting machine), RF (random forest) and MLP (multilayer perceptron) applied on predicting the dielectric properties and expressed as root mean square error (RMSE).

Machine learning method	*f* (MHz)	RMSE (Training set)			RMSE (Test set)		
		ε'	ε''	d_p (mm)	ε'	ε''	d_p (mm)
GBM	915	0.118	0.086	0.034	0.300	0.351	1.27
	2450	0.021	0.012	0.015	0.321	0.293	0.215
RF	915	0.506	0.611	2.06	0.652	0.655	2.15
	2450	0.404	0.280	0.456	0.782	0.335	0.558
MLP	915	0.245	0.130	0.573	0.239	0.166	0.692
	2450	0.192	0.137	0.146	0.224	0.147	0.142

result for 915 and 2450 MHz used 15 and 8 nodes at most for each tree, respectively. The finest number of trees found were 949 (ε'), 999 (ε'') and 997 (d_p) for 915 MHz, and 1140 (ε'), 1156 (ε'') and 1145 (d_p) for 2450 MHz. Table 5 presents the machine learning results for predicting the dielectric properties given the root mean square error as the performance criterion. The GBM model showed the lowest RMSE for the training set among all models (0.012–0.118), but for the test set, the values were quite higher. For 915 MHz, RMSE was 0.300 (ε'), 0.351 (ε'') and 1.27 (d_p), but a slight decrease is observed for 2450 MHz, 0.321 (ε'), 0.293 (ε'') and 0.215 (d_p). This might not only be related to using a different dataset. According to [53], a higher number of trees and a lower learning rate yield better performances at a higher computational cost.

Likewise, the RF model also relies on tuning hyper-parameters to attempt a better-performing model [57]. In this study, the maximum number of trees was 2000, while the number of variables selected at each node and the node size ranged from 1 to 7 for both. Different maximum numbers of levels (3–10) and nodes in each tree (5–50) were evaluated. The finest performance was obtained considering 7 variables for each node split, a maximum of 50 nodes in each tree and a node size equivalent to 1. The optimal number of levels was 15 and 30, respectively for 915 and 2450 MHz. For 915 MHz, the optimal number of trees was 160, 400 and 120 trees for ε', ε'' and d_p, respectively. Meanwhile, numbers of trees equal to 330 (ε'), 500 (ε'') and 240 (d_p) were obtained for 2450 MHz. The RF model showed the worst agreement with experimental data for the training set (0.280–2.06) and test set (0.335–2.15) among the models. The RMSE values of 0.652 (ε'), 0.655 (ε'') and 2.15 (d_p) for 915 MHz and 0.782 (ε'), 0.335 (ε'') and 0.558 (d_p) for 2450 MHz were found for the test dataset. Although GBM and RF are both based on the DT algorithm, the gradient algorithm seemed to have improved the performance of the GBM model concerning RF.

MLPs, widely employed supervised learning neural networks for prediction and classification, were applied with 11 input neurons and 1 output neuron for ε', ε'', or d_p. In the training phase, diverse hidden neuron quantities (1–30) and activation functions (tangent hyperbolic, logistic sigmoid, exponential, and identity) were tested for performance evaluation. The optimal configurations, determined by the training set, yielded the best results at 915 MHz with 4 hidden neurons for ε' (RMSE = 0.239), and 6 hidden neurons for ε'' (RMSE = 0.166) and d_p (RMSE = 0.692). At 2450 MHz, optimal results were achieved with 7 hidden neurons for ε' (RMSE = 0.224), 6 hidden neurons for ε'' (RMSE = 0.147), and 8 hidden neurons for d_p (RMSE = 0.142). Output activation functions identity and exponential were identified as optimal, while tangent hyperbolic and exponential served as ideal hidden activation functions. Notably, these configurations align with findings that demonstrated success in predicting model juice solutions at different sugar contents, temperatures, and field frequencies [54]. It is worth recognizing the significance of the training set in configuring the neural network, influencing hidden neurons, and determining optimal activation functions. The test set complements this process, independently assessing the model's generalization capacity to new data and ensuring robust predictions beyond the training.

Fig. 7 presents the variable importance plots for ε', ε'', and d_p. Notably, when considering relative importance higher than 50 %, the dielectric constant showcased temperature and electrical conductivity as the foremost predictors for GBM and RF. Conversely, MLP identified *a** and temperature as the primary variables at 915 MHz, while temperature, TS, and *a** held the highest importance at 2450 MHz. Examining the dielectric loss factor at 915 MHz, electrical conductivity emerged as the predominant factor, holding 100 % relative importance in all models. Additionally, in GBM, *a** and TTA claimed the second and third positions, respectively. RF assigned ashes and *a** to the second and third positions, while MLP allocated *b** and *L** to these respective positions. At 2450 MHz, all machine learning methods consistently identified electrical conductivity, temperature, and pH as the three most important variables influencing the dielectric loss factor. In terms of penetration depth, GBM highlighted electrical conductivity as the primary variable, followed in descending order by ashes, TTA, *b**, *L**, and

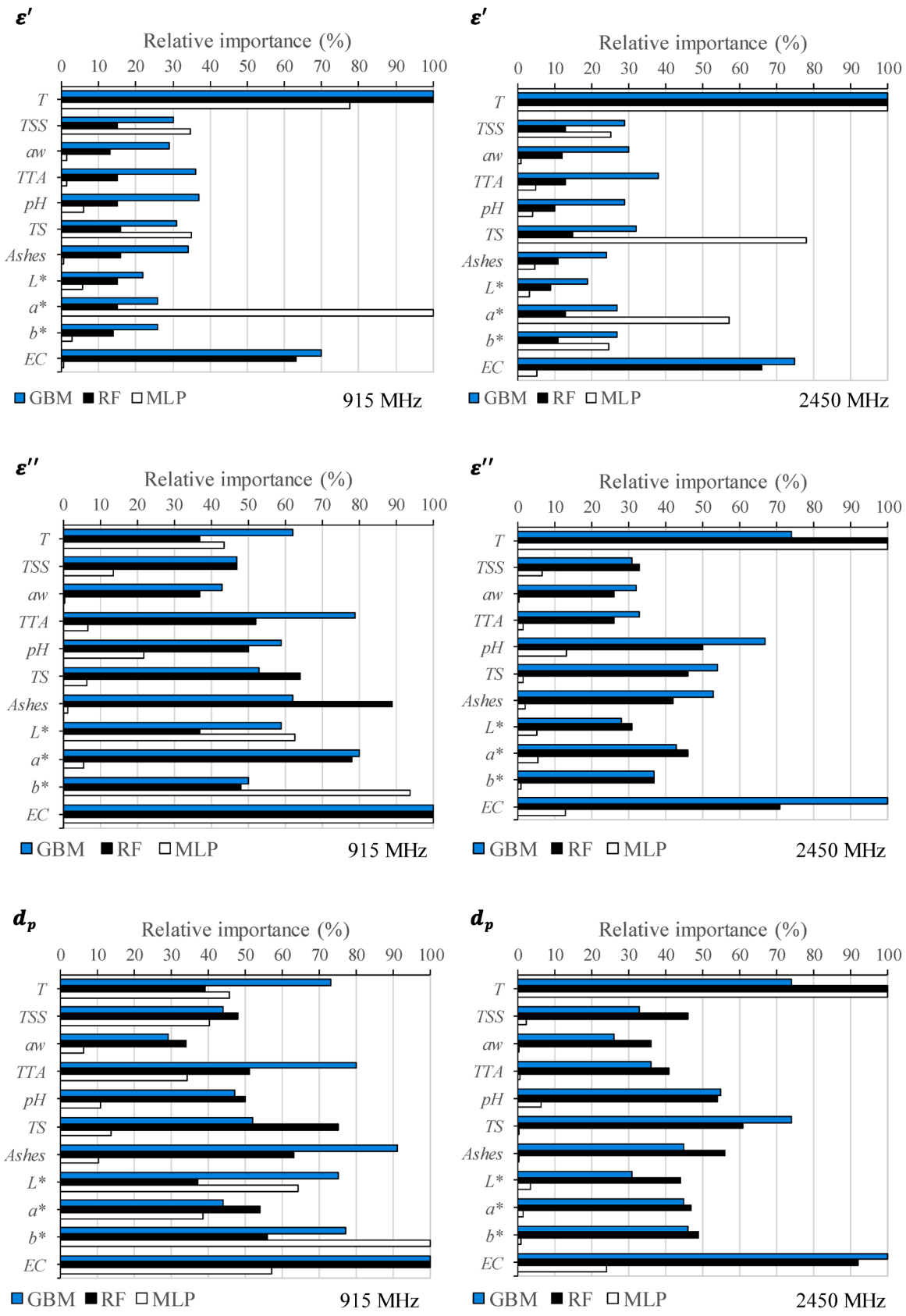


Fig. 7. Variable importance plot of dielectric constant (ε′), dielectric loss factor (ε″), and penetration depth (d_p) at 915 and 2450 MHz using GBM, RF and MLP models.

temperature. RF also emphasized electrical conductivity, with TS, ashes, b^*, a^*, and TTA following suit. MLP identified b^*, L^*, and electrical conductivity as the top three influential variables. Further scrutinizing the importance rankings at 2450 MHz, GBM prioritized electrical conductivity, while RF and MLP favored temperature. GBM underscored the significance of total solids (TS), temperature, and pH, each exceeding 50 % relative importance. RF, on the other hand, considered electrical conductivity, TS, ashes, and pH as pivotal factors. For MLP, temperature was the sole variable with a relative importance higher than 50 %, with electrical conductivity and pH securing positions of 24.1 % and 6.3 %, respectively, in the second and third slots.

In summary, this investigation revealed that temperature and electrical conductivity consistently emerged as pivotal predictors, demonstrating their robust ranking in all models. Notably, the variable importance plots underscored the substantial impact of ashes, pH, and total solids in certain scenarios, highlighting the influence of dilution, minerals, and acidity on dielectric properties. Moreover, the study observed a significant influence of color parameters, indicative of compounds, such as lycopene in guava and carotenoids in orange. A refined analysis, as detailed in the variable importance plots, provided a nuanced understanding of the intricate multivariate factors influencing the prediction of dielectric properties. While MLP exhibited notable strengths in capturing specific relationships, the choice of the optimal method may hinge on the specific objectives and nuances inherent in the dataset.

4. Conclusions

All machine learning techniques, notably multilayer perceptron (MLP), gradient boosting machine (GBM) and random forest (RF) excelled in predicting the temperature-dependent dielectric properties of the fruit juices based on physicochemical measurements. While lacking explicit mathematical or physical relationships compared to multiple linear regression (MLR), these models offer advantages such as effective handling of mixed datasets and suitability for non-linear systems. Key predictors included temperature and electrical conductivity, with composition and physicochemical attributes, such as total solids, acidity, pH, and color parameters also playing important roles. In summary, chemical analyses coupled with machine learning tools provided a swift and objective method for correlating and predicting the dielectric properties of this group of eight fruit juices. This approach, with its rapid learning and high accuracy, holds promise for advancing food processing simulation, particularly in microwave heating systems. Future studies can: (1) broaden the range of beverages in the group, (2) add new predictors to the data sets, or (3) make practical use of the adjusted machine learning tool, by using the predicted dielectric properties of a given product and temperature range, for the multi-physics simulation of a microwave applicator chamber.

CRediT authorship contribution statement

Rodrigo Nunes Cavalcanti: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis. **Vitor Pereira Barbosa:** Investigation, Formal analysis. **Jorge Andrey Wilhelms Gut:** Methodology, Writing – review & editing, Supervision. **Carmen Cecilia Tadini:** Writing – review & editing, Supervision, Resources, Funding acquisition, Conceptualization.

Declaration of competing interest

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

Acknowledgments

The authors acknowledge the financial supports from FAPESP - São Paulo Research Foundation (grants 2012/04073–0, 2013/07914–8, 2015/25697–0 and 2017/09730–2), CNPq - National Council for Scientific and Technological Development (grants 309548/2021–7, 316388/2021–1 and PIBIC-USP), and CAPES - Coordination for the Improvement of Higher Education Personnel (Finance Code 001).

References

[1] T. Lan, J. Wang, S. Bao, Q. Zhao, X. Sun, Y. Fang, et al., Effects and impacts of technical processing units on the nutrients and functional components of fruit and vegetable juice, Food Res. Int. 168 (2023) 112784, https://doi.org/10.1016/j.foodres.2023.112784.
[2] G.B. Awuah, H.S. Ramaswamy, A. Economides, Thermal processing and quality: principles and overview, Chem. Eng. Process. Process Intensif. 46 (2007) 584–602, https://doi.org/10.1016/j.cep.2006.08.004.
[3] M. Lin, H.S. Ramaswamy, Evaluation of phosphatase inactivation kinetics in milk under continuous flow microwave and conventional heating conditions, Int. J. Food Prop. 14 (2011) 110–123, https://doi.org/10.1080/10942910903147841.
[4] J. Tang, Unlocking potentials of microwaves for food safety and quality, J. Food Sci. 80 (2015) E1776–E1793, https://doi.org/10.1111/1750-3841.12959.
[5] T.K. Oishi, E.V.S. Pouzada, J.A.W. Gut, Experimental validation of a multiphysics model for the microwave-assisted pasteurization of apple juice, Digit. Chem. Eng. 5 (2022) 100053, https://doi.org/10.1016/j.dche.2022.100053.
[6] A.D. González-Monroy, G. Rodríguez-Hernández, C. Ozuna, M.E. Sosa-Morales, Microwave-assisted pasteurization of beverages (tamarind and green) and their quality during refrigerated storage, Innov. Food Sci. Emerg. Technol. 49 (2018) 51–57, https://doi.org/10.1016/j.ifset.2018.07.016.
[7] K.C. Amaro, G. Russo, D.L. Fan, J.A.W. Gut, C.C. Tadini, Modeling and experimental validation of the time-temperature profile, pectin methylesterase inactivation, and ascorbic acid degradation during the continuous flow microwave-assisted pasteurization of orange juice, Food Bioprod. Process. 144 (2024) 191–202, https://doi.org/10.1016/j.fbp.2024.01.015.
[8] É.S. Siguemoto, E. Purgatto, N.M.A. Hassimotto, J.A.W. Gut, Comparative evaluation of flavour and nutritional quality after conventional and microwave-assisted pasteurization of cloudy apple juice, LWT 111 (2019) 853–860, https://doi.org/10.1016/j.lwt.2019.05.111.
[9] Chang X., Zhang L., Xu Q., Zheng Z., Wang R., Li Z. Continuous flow microwave heating and sterilization for liquid food 2022;18:717–35. 10.1515/ijfe-2022-0130.
[10] R. Vadivambal, D.S. Jayas, Non-uniform temperature distribution during microwave heating of food materials-a review, Food Bioprocess Technol. 3 (2010) 161–171, https://doi.org/10.1007/s11947-008-0136-0.
[11] Y. Huang, J. Sheng, F. Yang, Q. Hu, Effect of enzyme inactivation by microwave and oven heating on preservation quality of green tea, J. Food Eng. 78 (2007) 687–692, https://doi.org/10.1016/j.jfoodeng.2005.11.007.
[12] C.P.C. Martins, R.N. Cavalcanti, S.M. Couto, J. Moraes, E.A. Esmerino, M.C. Silva, et al., Microwave processing: current background and effects on the physicochemical and microbiological aspects of dairy products, Compr. Rev. Food Sci. Food Saf. 18 (2019) 67–83, https://doi.org/10.1111/1541-4337.12409.
[13] P. Kumar, P. Coronel, J. Simunovic, V.D. Truong, K.P. Sandeep, Measurement of dielectric properties of pumpable food materials under static and continuous flow conditions, J. Food Sci. 72 (2007) 177–183, https://doi.org/10.1111/j.1750-3841.2007.00315.x.
[14] M.E. Sosa-Morales, L. Valerio-Junco, A. López-Malo, H.S. García, Dielectric properties of foods: reported data in the 21st century and their potential applications, LWT Food Sci. Technol. 43 (2010) 1169–1179, https://doi.org/10.1016/j.lwt.2010.03.017.
[15] A. Datta, G. Sumnu, V. Raghavan, Dielectric properties of foods, Eng. Prop. Foods (2014) 501–565, https://doi.org/10.1201/b16897-15.
[16] X. Zhu, W. Guo, Y. Jia, F. Kang, Dielectric properties of raw milk as functions of protein content and temperature, Food Bioprocess Technol. 8 (2015) 670–680, https://doi.org/10.1007/s11947-014-1440-5.
[17] J.G. Lyng, J.M. Arimi, M. Scully, F. Marra, The influence of compositional changes in reconstituted potato flakes on thermal and dielectric properties and temperatures following microwave heating, J. Food Eng. 124 (2014) 133–142, https://doi.org/10.1016/j.jfoodeng.2013.09.032.
[18] D. Salvi, D. Boldor, G.M. Aita, C.M. Sabliov, COMSOL Multiphysics model for continuous flow microwave heating of liquids, J. Food Eng. 104 (2011) 422–429, https://doi.org/10.1016/j.jfoodeng.2011.01.005.
[19] Y. Tao, Y. Tao, H. Yang, B. Yan, N. Zhang, Y. Zhang, et al., Microwave-induced thermal response and protein variation of oil-water biphasic systems in foods: a case study of anhydrous butter and skim milk, J. Food Eng. 372 (2024) 111997, https://doi.org/10.1016/j.jfoodeng.2024.111997.
[20] L. Xu, J. Su, H. Chen, J. Ye, K. Qin, W. Zhang, et al., Continuous-flow microwave heating system with high efficiency and uniformity for liquid food, Innov. Food Sci. Emerg. Technol. 91 (2024) 103556, https://doi.org/10.1016/j.ifset.2023.103556.
[21] Y. Zhang, C. Liu, X. Zheng, X. Zhao, L. Shen, M. Gao, Analysis of microwave heating uniformity in berry puree: from electromagnetic-wave dissipation to heat and mass transfer, Innov. Food Sci. Emerg. Technol. 90 (2023) 103509, https://doi.org/10.1016/j.ifset.2023.103509.

[22] M.T.K.T.K. Kubo, S. Curet, P.E.D. Augusto, L. Boillereaux, Artificial neural network for prediction of dielectric properties relevant to microwave processing of fruit juice, J. Food Process. Eng. 41 (2018) 1–16, https://doi.org/10.1111/jfpe.12815.
[23] H. Yang, B. Yan, W. Chen, D. Fan, Prediction and innovation of sustainable continuous flow microwave processing based on numerical simulations: a systematic review, Renew. Sustain. Energy Rev. 175 (2023) 113183, https://doi.org/10.1016/j.rser.2023.113183.
[24] H.W. Yang, S. Gunasekaran, Comparison of temperature distribution in model food cylinders based on Maxwell's equations and Lambert's law during pulsed microwave heating, J. Food Eng. 64 (2004) 445–453, https://doi.org/10.1016/j.jfoodeng.2003.08.016.
[25] R.N. Cavalcanti, C.F. Balthazar, E.A. Esmerino, C. Silva, R.S.L.L. Raices, J.A.W. W. Gut, et al., Correlation between the dielectric properties and the physicochemical characteristics and proximate composition of whole, semi-skimmed and skimmed sheep milk using chemometric tools, Int. Dairy J. 97 (2019) 120–130, https://doi.org/10.1016/j.idairyj.2019.05.018.
[26] X. Zhu, W. Guo, Y. Jia, Temperature-dependent dielectric properties of raw cow's and goat's milk from 10 to 4500MHz relevant to radio-frequency and microwave pasteurization process, Food Bioprocess Technol. 7 (2014) 1830–1839, https://doi.org/10.1007/s11947-014-1255-4.
[27] X. Zhu, W. Guo, Z. Liang, Determination of the fat content in cow's milk based on dielectric properties, Food Bioprocess Technol. 8 (2015) 1485–1494, https://doi.org/10.1007/s11947-015-1508-x.
[28] X. Zhu, W. Guo, X. Wu, Frequency- and temperature-dependent dielectric properties of fruit juices associated with pasteurization by dielectric heating, J. Food Eng. 109 (2012) 258–266, https://doi.org/10.1016/j.jfoodeng.2011.10.005.
[29] A.P. Franco, C.C. Tadini, J.A.W. Gut, Predicting the dielectric behavior of orange and other citrus fruit juices at 915 and 2450MHz, Int. J. Food Prop. 20 (2017) 1468–1488, https://doi.org/10.1080/10942912.2017.1347674.
[30] A.P. Franco, L.Y. Yamamoto, C.C. Tadini, J.A.W. Gut, Dielectric properties of green coconut water relevant to microwave processing: effect of temperature and field frequency, J. Food Eng. 155 (2015) 69–78, https://doi.org/10.1016/j.jfoodeng.2015.01.011.
[31] J. Peng, J. Tang, Y. Jiao, S.G. Bohnet, D.M. Barrett, Dielectric properties of tomatoes assisting in the development of microwave pasteurization and sterilization processes, LWT Food Sci. Technol. 54 (2013) 367–376, https://doi.org/10.1016/j.lwt.2013.07.006.
[32] S.O. Nelson, Dielectric spectroscopy of fresh fruit and vegetable tissues from 10 to 1800MHz, J. Microw. Power. Electromagn. Energy 40 (2005) 31–47, https://doi.org/10.1080/08327823.2005.11688523.
[33] T.A. Brinley, V.D. Truong, P. Coronel, J. Simunovic, K.P. Sandeep, Dielectric properties of sweet potato purees at 915MHz as affected by temperature and chemical composition*, Int. J. Food Prop. 11 (2008) 158–172, https://doi.org/10.1080/10942910701284291.
[34] P. Yaghmaee, T.D. Durance, Predictive equations for dielectric properties of NaCl, D-sorbitol and sucrose solutions and surimi at 2450MHz, J. Food Sci. 67 (2002) 2207–2211, https://doi.org/10.1111/j.1365-2621.2002.tb09528.x.
[35] C.D. Everard, C.C. Fagan, C.P. O'Donnell, D.J. O'Callaghan, J.G Lyng, Dielectric properties of process cheese from 0.3 to 3GHz, J. Food Eng. 75 (2006) 415–422, https://doi.org/10.1016/j.jfoodeng.2005.04.027.
[36] M. Zhao, G. Downey, C.P O'Donnell, Exploration of microwave dielectric and near infrared spectroscopy with multivariate data analysis for fat content determination in ground beef, Food Control 68 (2016) 260–270, https://doi.org/10.1016/j.foodcont.2016.03.031.
[37] M. Kent, R. Knöchel, F. Daschner, U.K. Berger, Composition of foods including added water using microwave dielectric spectra, Food Control 12 (2001) 467–482, https://doi.org/10.1016/S0956-7135(01)00021-4.
[38] H. Lizhi, K. Toyoda, I. Ihara, Discrimination of olive oil adulterated with vegetable oils using dielectric spectroscopy, J. Food Eng. 96 (2010) 167–171, https://doi.org/10.1016/j.jfoodeng.2009.06.045.
[39] L. Vidal, L. Antúnez, A. Rodríguez-haralambides, A. Giménez, K. Medina, E. Boido, et al., Relationship between astringency and phenolic composition of commercial Uruguayan Tannat wines : application of boosted regression trees, Food Res. Int. 112 (2018) 25–37, https://doi.org/10.1016/j.foodres.2018.06.024.
[40] Codex Alimentarius Commission. (CXS 247-2005). General Standard for Fruit Juices and Nectars. 2005.
[41] Food and Drug Administration. (21 CFR 101.30). Percentage juice declaration for foods purporting to be beverages that contain fruit or vegetable juice. 2024.
[42] Williams S. Official methods of analysis of the Association of Official Analytical Chemists. 14th ed. Arlington (Va.) : Association of Official Analytical Chemists; 1984.
[43] D. Kimball, Brix and soluble solids, Kimball D. Citrus Processing: Quality Control and Technology, Springer Netherlands, Dordrecht, 1991, pp. 7–33, https://doi.org/10.1007/978-94-011-3700-3_2.
[44] W. Routray, V. Orsat, Recent advances in dielectric properties–measurements and importance, Curr. Opin. Food Sci. 23 (2018) 120–126, https://doi.org/10.1016/j.cofs.2018.10.001.
[45] A. Abea, P. Gou, M.D. Guàrdia, P. Picouet, M. Kravets, S. Bañón, et al., Dielectric heating: a review of liquid foods processing applications, Food Rev. Int. 39 (2023) 5684–5702, https://doi.org/10.1080/87559129.2022.2092746.
[46] E.S. Siguemoto, J.A.W. Gut, Dielectric properties of cloudy apple juices relevant to microwave pasteurization, Food Bioprocess Technol. 9 (2016) 1345–1357, https://doi.org/10.1007/s11947-016-1723-0.
[47] P.O. Risman, Terminology and notation of microwave power and electromagnetic energy.pdf, J. Microwave Power Electromagn. Energy 26 (1991) 243–250.
[48] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32, https://doi.org/10.1023/A:1010933404324.
[49] Qiu S., Wang J., Tang C., Du D. Comparison of ELM, RF, and SVM on E-nose and E-tongue to trace the quality status of mandarin (Citrus unshiu Marc .) 2015;166: 193–203. 10.1016/j.jfoodeng.2015.06.007.
[50] E. Vigneau, P. Courcoux, R. Symoneaux, L. Guérin, A. Villière, Random forests : a machine learning methodology to highlight the volatile organic compounds involved in olfactory perception, Food Qual. Prefer. 68 (2018) 135–145, https://doi.org/10.1016/j.foodqual.2018.02.008.
[51] D. Zhang, Y. Guo, S. Rutherford, C. Qi, X. Wang, P. Wang, et al., The relationship between meteorological factors and mumps based on Boosted regression tree model, Sci. Total Environ. 695 (2019) 133758, https://doi.org/10.1016/j.scitotenv.2019.133758.
[52] Y. Song, H. Zhou, P. Wang, M. Yang, Prediction of clathrate hydrate phase equilibria using gradient boosted regression trees and deep neural networks, J. Chem. Thermodyn. 135 (2019) 86–96, https://doi.org/10.1016/j.jct.2019.03.030.
[53] Elith J., Leathwick J.R., Hastie T. A working guide to boosted regression trees 2008:802–13. 10.1111/j.1365-2656.2008.01390.x.
[54] M.T.K. Kubo, S. Curet, P.E.D. Augusto, L. Boillereaux, Artificial neural network for prediction of dielectric properties relevant to microwave processing of fruit juice, J. Food Process. Eng. 41 (2018) 1–16, https://doi.org/10.1111/jfpe.12815.
[55] F. Icier, C. Ilicali, Temperature dependent electrical conductivities of fruit purees during ohmic heating, Food Res. Int. 38 (2005) 1135–1142, https://doi.org/10.1016/j.foodres.2005.04.003.
[56] Cavalcanti R.N., Balthazar C.F., Esmerino E.A., Silva C., Raices R.S.L., Gut J.A.W., et al. Correlation between the dielectric properties and the physicochemical characteristics and proximate composition of whole, semi-skimmed and skimmed sheep milk using chemometric tools 2019;97:120–30. 10.1016/j.idairyj.2019.05.018.
[57] E. Vigneau, P. Courcoux, R. Symoneaux, L. Guérin, A. Villière, Random forests: a machine learning methodology to highlight the volatile organic compounds involved in olfactory perception, Food Qual. Prefer. 68 (2018) 135–145, https://doi.org/10.1016/j.foodqual.2018.02.008.