2022 19th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE). Mexico City, Mexico. November 9-11, 2022

Microwave dielectric spectroscopy for determination of ethanol concentration in brandy

Edel-Serafin Hernandez-Gomez
Department of Electronic
National Institute of Optical Astrophysics and Electronics
Puebla, Mexico
0000-0002-9724-1542

Jose-Luis Olvera-Cervantes
Department of Electronic
National Institute of Optical Astrophysics and Electronics
Puebla, Mexico
jolvera@inaoep.mx

Nefi-David Pava-Chipol
Campus Acayucan
Technological Institute of Mexico
Acayucan, Mexico
sowbran@gmail.com

Abstract— **Brandy is one of the most consumed alcoholic drinks in Mexico. A study by the National Laboratory for Consumer Protection of alcoholic drinks marketed in Mexico, priced at no more than $ 130.00 mexican pesos, found that while most of the brandy-type drinks analyzed meet the standards, some do not meet the ethanol content that they declare. For this reason, it is necessary to determine ethanol content in alcoholic drinks to assure the quality and authenticity. This is an important opportunity for microwave dielectric spectroscopy which can be a faster and cheaper technique for the determination of ethanol in alcoholic drinks that other techniques such as chromatography. In this work, the complex relative permittivity of liquid samples with different ethanol concentrations in brandy is presented. The brandy considered was Presidente® because it complies with the Mexican Official Standard. The study covers a concentration range from 0 to 50 % v/v. The measurements are performed with an open coaxial probe in a range from 500 MHz to 20 GHz and the Havriliak-Negami model parameters are presented for the different samples. A data analysis process was performed to predict the ethanol concentration by means of these dielectric parameters. The model was created with the training set and evaluated with the testing set. The feature chosen for the model was the dielectric constant under direct current. The mean squared error resulting was 1.071 while the model found in the literature and which is based on relaxation time was 1.21. The information presented in this work could be employed for the design and implementation of microwave microstrip sensors which are able to measure ethanol content in brandy in real-time and cheaply.**

Keywords— permittivity, brandy, sensor, model, ethanol.

I. INTRODUCTION

Beverages can be defined as “any liquid that is consumed by drinking”. It is composed of different groups of food products which include fruit drinks, synthetic drinks, tea, alcoholic drinks, milk drinks, coffee, chocolate drinks, etc [1]. Beverages can be classified in alcoholic and nonalcoholic drinks. Beers, wines, liqueurs, and others belong to alcoholic drinks. In similar manner, nonalcoholic drinks are grouped in carbonated drinks, noncarbonated drinks, and hot beverages. Juices, energy drinks, carbonated drinks, tea, coffee, and bottled water are included in this category [1].

Alcoholic drinks are popular products that are consumed worldwide for different purposes, such as entertainment, customs, religious reasons, among others. These are produced by distillation and fermentation. Of this set, those produced on an industrial scale for sale on the market must comply with regulatory standards. Ethanol is one of its main components, an important indicator of quality and is measured as alcohol content (abv) [2]. It is important to mention that brandy is a drink that is made from the distillation of wines from 100% grape must, whose distillate must undergo a maturation process in oak, white oak or oak wood containers for a minimum period of 6 months. Its alcoholic content is 35 to 55% [3]. In Mexico, according to the National Survey on the Consumption of Drugs, Alcohol and Tobacco (ENCODAT) 2016-2017, brandy ranked second among the most consumed alcoholic drinks, only after beer [4]. It is worth noting that alcoholic drinks can be adulterated with ethanol. According to samples acquired by the Federal Commission for Protection Against Sanitary Risks (COFEPRIS) from Mexico, ethanol has been detected in various adulterated beverages, which even in small doses higher that those reported in its labels, can cause damage to the body and cause blindness or even, in extreme cases, death [5].

As mentioned before, the content of ethanol in alcoholic drinks is determined to assure the quality. It is used for authenticity of drink products as well. Lately, the ethanol content in alcoholic drinks has been measured by several techniques such as gas chromatography-mass spectrometry (GC-MS), H NMR spectrometry, thermal infrared enthalpimetry, Near-Infrared Spectrophotometry, electrochemical enzyme biosensors, digital image-based method, membraneless gas-liquid separation flow system, flatbed scanner and automated digital image analysis, propagating waves, reversible solid sensor and colorimetric oscillating signals (COS) [6]. There are a few works about microwave dielectric spectroscopy for the determination of ethanol in alcoholic drinks which can be a fast and cheap technique. It can be used to track non-destructive alcoholic fermentation process [7].

Regarding microwave dielectric spectroscopy, models aimed to ethanol content determination using dielectric relaxation parameters have been established by means of simple linear regression [7], [8]. The evaluation of the obtained models was only carried out with training samples (samples used for the creation of the model) and test samples were not used (samples not included in the creation of the model). Evaluation with test samples provides an objective result of model performance.

In this paper, the study of the complex relative permittivity of the Presidente® product is presented because it is a brandy-type distilled beverage which complies with the Mexican Official Standard NOM-142-SSA1/SCFI-2014, Alcoholic Beverages, Sanitary Specifications, Sanitary and Commercial Labeling according to the Federal Consumer Attorney (PROFECO) from Mexico. It is worth mentioning that there are other brands of brandy which are sold in the market and do not comply with the corresponding standard. Specifically, they do not comply with the amount of ethanol content, that

978-1-6654-5508-4/22/$31.00 ©2022 IEEE

is, they report a different ethanol content than that measured by PROFECO. For the purity analysis of the Presidente® drink considering the permittivity, different concentrations of ethanol in Presidente® ranging from 0 to 50 % v/v were considered. These concentrations were measured in the frequency range from 500 MHz to 20 GHz. This range was considered because it contains the 902-928 MHz, 2400-2500 MHz and 5725-5875 MHz bands designated for industrial, scientific and medical (ISM) applications [9]. For each sample, the dielectric relaxation parameters according to Havriliak-Negami model are presented. The objective of this work was to dielectrically characterize several liquid samples with different concentrations of ethanol in Presidente® in order to determine this concentration using a statistical technique that employs, as independent variables, dielectric relaxation parameters.

II. Dielectric Parameters

The terms dielectric constant (ε') and loss factor (ε'') define the complex relative permittivity (ε^*) to that of free space which is expressed in (1).

$$\varepsilon^* = \varepsilon' - \varepsilon'' \quad (1)$$

The dielectric constant is defined as the measure of the ability of a material to store electromagnetic energy [10]. The dielectric loss factor is characterized by the amount of electromagnetic energy converted into heat in a material [11].

Havriliak-Negami model is the most common model used to describe the electrical behavior of aqueous solutions or tissues. It is an empirical modification of the Debye relaxation model and represents the asymmetry and amplitude of the dielectric dispersion curve:

$$\varepsilon^* = \varepsilon_\infty + \frac{\varepsilon_s - \varepsilon_\infty}{\left[1+(j\omega\tau)^{(1-\alpha)}\right]^\beta} - \frac{j\sigma_{dc}}{\omega\varepsilon_0} \quad (2)$$

ε_s and ε_∞ are the dielectric constants under DC and at infinity frequency, respectively. σ_{dc} is the DC conductivity. τ is the characteristic relaxation time of the medium, which is the time required for dipoles to become oriented in the presence of an electric field or the time needed to disorient the dipoles after the electric field is removed. α and β are empirical variables that account for the distribution of the relaxation time and the asymmetry of the relaxation time distribution, respectively [12], [13]. The relaxation frequency indicates the frequency in which a group of dipoles of material no longer follows to the electrical field and in which the loss factor reaches a maximum. This frequency occurs at $\omega = 1 / \tau$ [14], [15].

III. Methodology

A. Experimental matrix

Table 1 shows the composition of the samples of ethanol (9014-02, J. T. Baker, USA) in Presidente® considered in this work with their respective denominations. The concentration level proposed goes from 0 to 50 % v/v.

B. Measurement of dielectric properties

A wide-frequency range open-coaxial probe was used to determine the permittivity of liquid samples [16]. The set-up consisted of a vector network analyzer (Keysight Technologies, N9918A, USA), a 50 mL glass container, an open coaxial probe (Keysight Technologies, Slim Form Probe 030, USA) and a 50Ω flexible cable. The Agilent-85070E coaxial probe kit consisting of an open circuit, a short circuit and 30 mL of distilled-deionized water at room temperature (25 °C) was employed to get the system calibrated. For both measurement and calibration, the probe tip was immersed 2 cm inside the samples, resulting in a distance of about λ/8 from the probe tip to the bottom of the sample at the lowest frequency [16]. The samples were measured by quintuplicate. The experimental set up is shown in Fig. 1.

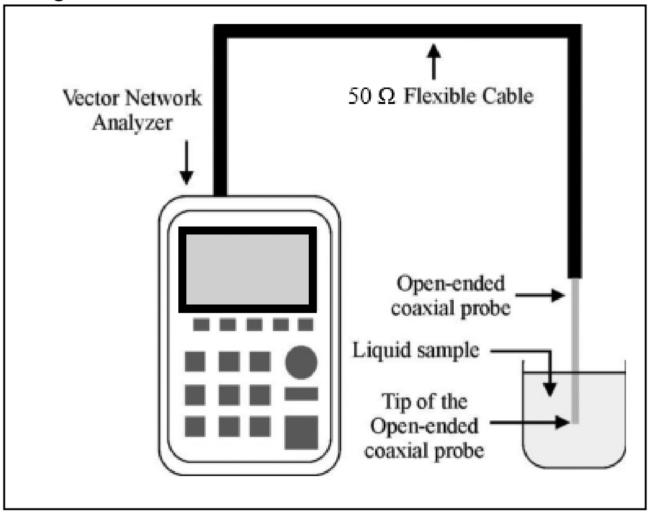

Figure 1. Experimental set up

TABLE I CONCENTRATION OF ETHANOL IN PRESIDENTE®

Liquid sample ID	Ethanol (mL)	Presidente® (mL)	Concentration of ethanol (% v/v)
E1	0.00	40.00	0.00
E2	0.62	39.38	1.56
E3	1.25	38.75	3.12
E4	2.50	37.50	6.25
E5	5.00	35.00	12.50
E6	6.00	34.00	15.00
E7	8.00	32.00	20.00
E8	9.00	31.00	22.50
E9	10.00	30.00	25.00
E10	11.00	29.00	27.50
E11	12.00	28.00	30.00
E12	13.00	27.00	32.50
E13	14.00	26.00	35.00
E14	15.00	25.00	37.50
E15	16.00	24.00	40.00
E16	17.00	23.00	42.50
E17	18.00	22.00	45.00
E18	19.00	21.00	47.50
E19	20.00	20.00	50.00

C. Data analysis process

Data analysis is the collection, transformation, and organization of data in order to draw conclusions, make predictions, and drive informed decision-making [17]. In this work, data ingestion, data cleaning, model training and candidate model evaluation were performed. In data ingestion, information was obtained to create the final model. Anomaly detection and eliminating duplicate were considered in data cleaning. Model training was carried out finding relationships between variables by correlation matrix, reducing the quantity of variables with backward elimination in multiple linear regression and making a data segregation. Finally, the model was created with training samples. In candidate model evaluation, the model was evaluated with testing samples. The data analysis process in shown in Fig. 2. It is important to mention that the data analysis process was carried out with Python (Python Software Foundation, DE, USA). The most important libraries used were: pandas, numpy, seaborn, statsmodels.api and sklearn.model_selection.

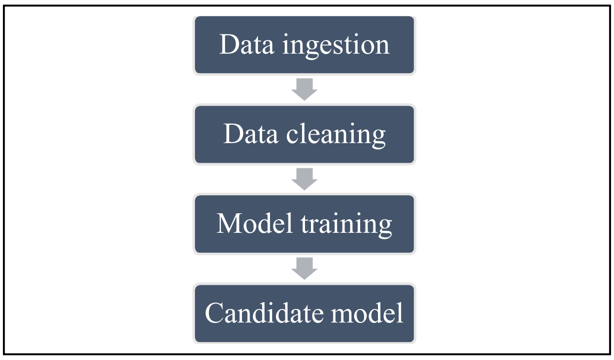

Figure 2. Data analysis process.

IV. Results and Discussion

A. Data ingestion

The measurements of the complex permittivity of the 19 liquid samples with different ethanol concentration considered in Table 1 were carried out. The frequency range is from 500 MHz to 20 GHz with 1001 points. For each of the samples, the parameters of the Havriliak-Negami model were extracted from (2) and are shown in Table 2.

TABLE II HAVRILIAK-NEGAMI MODEL PARAMETERS OF THE LIQUID SAMPLES WITH DIFFERENT ETHANOL CONCENTRATION IN PRESIDENTE®

Liquid sample ID	Havriliak-Negami model parameters					
	ε_∞	ε_s	σ_{dc} (µS/cm)	α	β	τ (ps)
E1	8.00	62.50	4.00	0.00	1.00	27.70
E2	8.10	61.00	4.00	0.00	1.00	26.90
E3	8.10	60.50	4.00	0.00	1.00	27.70
E4	8.10	59.00	4.00	0.00	1.00	28.10
E5	7.80	57.00	4.00	0.00	1.00	31.00
E6	7.75	55.00	4.00	0.00	1.00	31.60
E7	7.00	47.50	4.00	0.00	1.00	40.70
E8	5.00	45.00	4.00	0.00	1.00	44.20
E9	7.00	52.00	4.00	0.00	1.00	35.90
E10	5.40	44.00	4.00	0.00	1.00	46.50
E11	6.90	48.00	4.00	0.00	1.00	40.70
E12	5.78	42.50	4.00	0.00	1.00	54.90
E13	5.20	41.42	4.00	0.00	1.00	54.90
E14	4.90	37.71	4.00	0.00	1.00	67.10
E15	6.00	40.00	4.00	0.00	1.00	68.20
E16	5.00	40.00	4.00	0.00	1.00	62.50
E17	5.00	40.50	4.00	0.00	1.00	62.00
E18	4.90	39.00	4.00	0.00	1.00	62.00
E19	5.60	35.50	4.00	0.00	1.00	73.80

It is important to mention that the parameters are obtained from the entire frequency range. The dielectric constant shows an adjustment of $R^2 \geq 0.90$ for all samples. On the other hand, the loss factor has an adjustment of $R^2 \geq 0.98$ for all samples. σ_{dc}, α and β are parameters that don't change, that is, they keep their values constant for all the samples. ε_∞ exhibits a change, but it doesn't show an apparent trend according with the ethanol concentration. ε_s and τ are parameters that decrease and increase, respectively, with the concentration. In fact, these behaviors have been observed in water−ethanol mixtures. The higher the ethanol concentration, the lower the ε_s and the higher the τ [18].

B. Data cleaning

Regarding data cleaning, an outlier detection was carried out, which was based on the Local Outlier Factor method, on the data shown in Table 1. In this, the local outlier of each sample was calculated to indicate the degree of similarity with neighboring samples. The sample or neighboring data is the one that the values of its variables are the closest to a specific data [19]. The samples identified as atypical were E7, E8, E9 and E11. Probably, these are outlier because they could have been manufactured incorrectly. Precisely, outlier detection has many applications such as manufacturing quality, which indicates what samples could be poorly manufactured [20]. In future works, not only several samples with the same concentration could be made for avoiding samples from error, but also the quantity of samples with different concentrations could increase. This could improve the performance of the model. Subsequently, a Duplicate Instance Removal was carried out. It wasn't found samples duplicated. At the end of data cleaning, samples E7, E8, E9 and E11 were discarded from the database.

C. Model training

In order to establish the relationships between features or independent variables, which in this case are dielectric relaxation parameters, with the ethanol content, a correlation matrix was performed. This is showed in Fig. 3. As it can see, the empty cells means that there is no correlation between variables. σ_{dc}, α and β don't have correlation with EC (ethanol concentration). On the other hand, ε_∞, ε_s and τ present the values -0.933, -0.981 and 0.968, respectively, as Pearson correlation coefficients respect with EC. These features were considered for multiple linear regression. Multiple linear regression models are a generalization of simple linear regression in cases where there are more than one independent, predictor variable or feature. The aim of such multiple regression is therefore to explore and quantify the relationship between a numerical dependent variable and one or more qualitative or qualitative predictor variables [21].

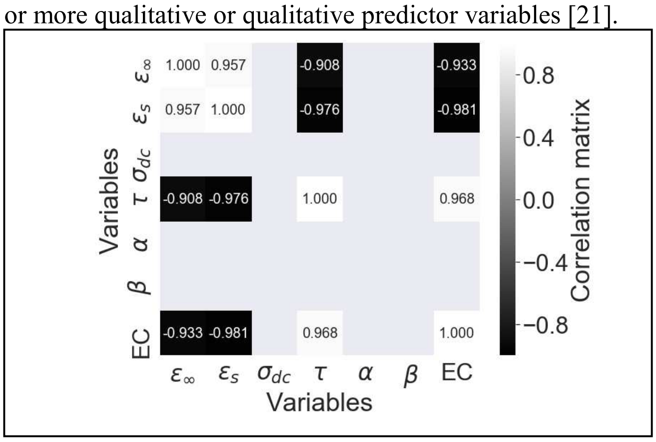

Figure 3. Pearson correlation coefficient matrix.

The backward elimination in multiple linear regression was implemented using p value and a significance level = 0.05. The feature ε_s was that which showed p value under significance level, that is, its Pearson correlation coefficient respect to ethanol content is significant different from 0. After feature selection, the dataset formed by 15 samples was divide into training and testing set. 80 % of dataset was used for training set and 20 %, for testing set. The linear regression produced the model which is showed in (3):

$$EC = 111.65 - 1.79\varepsilon_S \quad (3)$$

As it was mentioned before, this was formed by training set. The 95 % confidence interval of the interception (111.65) is from 98.56 to 124.75 and this interval for ε_S coefficient (-1.79) is from -2.06 to -1.52. The model can be observed in Fig. 4.

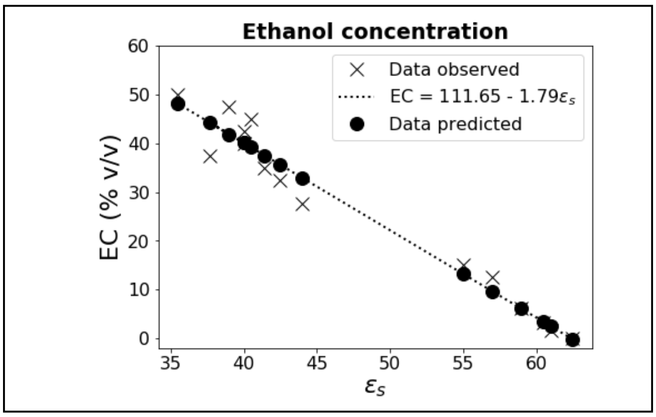

Figure 4. CE prediction model.

It is important to mention that there is no data in the range of 12.5 to 27.5 % v/v of ethanol concentration because the samples E7, E8 and E9 covered that interval and were discarded in data cleaning, however a sample of testing set with EC = 15 % v/v (E6) is considered for model performance evaluation.

D. *Candidate model evaluation*

The testing set is composed of the samples E4, E6 and E15. The evaluation results are showed in Table 3.

TABLE III MEAN SQUARED ERROR, ROOT MEAN SQUARED ERROR AND DETERMINATION COEFFICIENT OF MODEL EVALUATED

Mse	1.071
Rmse	1.035
R^2	0.998

The mean squared error (mse), root mean squared error (rmse) and determination coefficient (R^2) are common model evaluation metrics used for regression purposes [22]. In literature, there is other work that reports the model performance for determining alcohol content in brandy by means of permittivity. This showed a mse of 1.21. The feature considered was τ [7]. This means that the model presented in this work is better than that reported in literature considering the mse.

E. *Discussion*

In this work and as it was mentioned before, the feature used for determining alcohol content in brandy was ε_s. In this way, the dielectric properties of water greatly influence the dielectric properties of food and beverages, and alcoholic drinks are no exception. Strong negative correlations have also been recorded between the concentration of ethanol in water and ε_s. Reference [23] found a relationship between ethanol concentration in water in 1978. ε_s increased steadily with decreasing of ethanol content. References [15] and [24] reported the same behavior in 2017 and 2015, respectively.

ε_s is related with electric fields at frequencies where no relaxational effects occur. In this case, the delay between electric polarization in the liquid and the electric field from coaxial open probe is called relaxational effect. ε_s is a measure of the polarization character, and the measurements for liquids provide valuable information about the local structure due to the molecular orientation and interactions [25]. This is required to understand molecular interaction in liquids in the field of science and industry. The dielectric parameter analysis was done from 500 MHz to 20 GHz in both this work and literature because water-ethanol solution relaxation process can be characterized in this frequency range [15].

The information presented in this work could be employed for the design and implementation of microstrip sensors which are able to measure ethanol content in brandy in real-time. It is important to mention that microstrip technology is a cheap technique in order to create a microwave sensor and could replace the open coaxial probe which is more expensive.

V. CONCLUSION

In this work, the complex dielectric permittivity, measured with an open coaxial probe, is presented from 500 MHz to 20 GHz for several samples with different ethanol concentrations in brandy Presidente®. Data ingestion, data cleaning, model training and candidate model evaluation were performed for the data process. This showed that the ethanol concentration prediction is carried out considering the dielectric constant under direct current as the feature. The mean squared error of this model generated is lower than that of the model presented in literature and uses dielectric parameters as features as well. The results of this work could be used for the design of microwave microstrip sensors aimed to measure ethanol content cheap and fast.

ACKNOWLEDGMENT

The authors would like to thank Conacyt for the support of the project.

REFERENCES

[1] A. M. Grumezescu, A. M. Holban, "Biotechnological progress and beverage consumption", United Kingdom: Woodhead publishing. 2020.

[2] M. Rocchia, M. Ellena, G. Zeppa. Determination of Ethyl Alcohol Content in Red Wines with an Optical Alcood hol Meter Based on Nanostructured Silicon. J. Agric. Food Chem, vol 55, no. 15, pp. 5984-5993, 2007.

[3] SEGOB, "NORMA Oficial Mexicana NOM-199-SCFI-2017, Bebidas alcohólicas Denominación, especificaciones fisicoquímicas, información comercial y métodos de prueba", 2017. Accessed on: June

29, 2022. [Online]. Available: http://www.dof.gob.mx/nota_detalle.php?codigo=5502882&fecha=30/10/2017.

[4] SEGOB, "Encuesta Nacional de Consumo de Drogas, Alcohol y Tabaco, ENCODAT 2016-2017", 2017. Accessed on: June 29, 2022. [Online]. Available: https://www.gob.mx/salud%7Cconadic/acciones-y-programas/encuesta-nacional-de-consumo-de-drogas-alcohol-y-tabaco-encodat-2016-2017-136758

[5] SEGOB, "Boletín 094/2017", 2017. Accessed on: June 30, 2022. [Online]. Available: https://www.gob.mx/profeco/prensa/boletin-de-prensa-094-2017-profeco-y-el-gobierno-de-guerrero-firman-acuerdo-para-combatir-la-venta-de-alcohol-adulterado-en-destinos-turisticos?idiom=es-MX

[6] T. Somboon, R. Phatchana, W. Tongpoothorn, S. Sansuk, "A simple and green method for determination of ethanol in liquors by the conductivity measurement of the uncatalyzed esterification reaction", LWT, vol 154, no. 15, 2022.

[7] X. Bohigas, J. Tejada, "Dielectric characterization of alcoholic beverages and solutions of ethanol in water under microwave radiation in the 1–20 GHz range", Food Res. Int., vol 43, no. 6, pp. 1607-1613, 2010.

[8] T. Kaur-Kataria, M. E. Sosa-Morales, J. L. Olvera-Cervantes, A. Corona-Chavez. "Dielectric properties of tequila in the microwave frequency range (0.5–20 GHz) using coaxial probe", Int. J. Food Prop., vol. 20, pp. S377-S384. 2017.

[9] Federal Communications Commision, "FCC ONLINE TABLE OF FREQUENCY ALLOCATIONS", 2022. Accessed on: September 10, 2022. [Online]. Available: https://www.google.com.mx/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjP0r7W3Yr6AhVhLkQIHSlbCnAQFnoECAUQAQ&url=https%3A%2F%2Fwww.fcc.gov%2Ffile%2F21474%2Fdownload&usg=AOvVaw0xDgR2LYKoMRZczSTHoUM_

[10] M. A. Stuchly, S. S. Stuchly, "Dielectric Properties of Biological Substances —Tabulated," J Microw Power Electromagn Energy, vol 15, no. 1, pp. 19-25, 1980.

[11] S. O. Nelson, "Electrical properties of grain and other food materials". J. Food Process Pres., vol 2, no. 2, pp. 137-154, 1978.

[12] G. G. Raju. Dielectrics in Electric Fields. USA: Marcel Dekk, . 2003, pp.97–107.

[13] A. La-Gioia, E. Porter, I. Merunka, A. Shahzad, S. Salahuddin, M. Jones, M. O-Halloran, "Open-Ended Coaxial Probe Technique for Dielectric Measurement of Biological Tissues: Challenges and Common Practices," Diagnostic, vol 8, no. 2, 2018.

[14] H. Frohlich. "Theory of Dielectrics: Dielectric Constant and Dielectric Loss," UK: Oxford University Press, 1987, pp. 61–67.

[15] S. Mehrotra, A. Kumbharkhane, A. Chaudhari. "Binary Polar Liquids: Structural and Dynamic Characterization Using Spectroscopic Methods," USA: Elsevier, 2017, pp. 18–19.

[16] T. K. Kataria, M. E. Sosa-Morales, J. L. Olvera-Cervantes, A. Corona-Chávez. "Dielectric Properties of Guava, Mamey Sapote, Prickly Pears and Nopal in the Microwave Range", Int. J. Food Prop., vol 20, no. 12, pp. 2944–2953, 2017.

[17] S. Bhattacharyya, K. Mondal, K. B. Prakash, N. K. Mondal, J. P. Singh. "Cognitive Data Models for Sustainable Environment", India: Academic Press. 2022.

[18] H. Schubert, M. Regier. "The Microwave Processing of Foods", USA: Elsevier. 2005.

[19] D. Corrales-Camilo, J. C. Corrales, A. Ledezma. "How to Address the Data Quality Issues in Regression Models: A Guided Process for Data Cleaning". Simmetry, vol 10, no. 4, 2018.

[20] H. Belyadi, A. Haghighat. "Machine Learning Guide for Oil and Gas Using Python", USA: Gulf Professional Publishing. 2021.

[21] M. M. Rodríguez del Águilaa, N. Benítez-Parejo. "Simple linear and multivariate regression models", Basic statistics for busy clinicians (vi), vol 39, no. 31, pp. 59-173, 2011.

[22] A. Artasanchez, P. Joshi. "Artificial Intelligence with Python", United Kingdom: Packt Publishing. 2020.

[23] A. D'prano, D. I. Donato, E. Caponetti. "The Static Dielectric of solutions of water in n-alcohols at 15, 25, 35 and 45 °C", J. Solut. Chem. vol 8, no. 2, pp. 135-146, 1979.

[24] C. Wohlfarth. "Static Dielectric Constants of Pure Liquids and Binary Liquid Mixtures", Berlin: Springer. 2015.

[25] T. Moriyoshi, T. Ishii, Y. Tamai, M. Tad. "Static Dielectric Constants of Water 4- Ethanol and Water + 2-Methyl-2-propanol Mixtures from 0.1 to 300 MPa at 298.15 K", J. Chem. Eng., vol 35, pp. 17-20. 1990.