Original Article
Food Science and Technology

ISSN 0101-2061 (Print)
ISSN 1678-457X (Online)

DOI: https://doi.org/10.1590/fst.89222

Classification of chocolate according to its cocoa percentage by using Terahertz time-domain spectroscopy

Jimy Frank OBLITAS CRUZ[1]*

Abstract

Feasibility of a non-destructive classification of chocolate based on its cocoa content was examined by using a Terahertz time-domain spectroscopy system combined with a multivariate analysis. For this purpose, the spectra from 0.5 THz to 10 THz of 5 chocolate samples (50%, 60%, 70%, 80% and 90% of cocoa) were examined. The acquired data matrices were analyzed by using a Fourier Transform, obtaining the dielectric function and the absorbance curve. Based on the latter, samples were classified by using 24 models of mathematical classification, achieving differences of around 93% through the model of Gaussian SVM algorithm with a kernel scale of 0.35 and a one-against-one multiclass method. This was reduced by using a Main Component Analysis, obtaining most of the spectral variations with PC1 (63.8%) and PC2 (36.2%). It was concluded that the combined processing and classification of images obtained from Terahertz time-domain spectroscopy, as well as the use of machine learning algorithms, can be used to successfully classify chocolates with different percentages of cocoa.

Keywords: Terahertz spectroscopy; multivariate analysis; cocoa; chocolate.

Practical Application: Terahertz time-domain spectroscopy has been shown to have great potential as an emerging nondestructive tool for food inspection. The main applications of THz in food industries include detection of moisture, foreign bodies, inspection, and quality control.

1 Introduction

The so-called non-contact and non-destructive methods, such as NIR spectroscopy (Wang et al., 2022) and multispectral/hyperspectral images (De-la-Torre et al., 2020), images in the visible range (Castro et al., 2017), RAMAN spectroscopy (Xu et al., 2020), nuclear magnetic resonance (Balthazar et al., 2021), have been widely used in the food sector as they are sensitive to intra-molecular vibration (Wang et al., 2017) and have increasingly been applied as a powerful analytical tool for determining food quality, as well as for identifying the geographical origin. This boom in studies at this frequency is due to access to instruments that are available, but some spectra of the intermediate band or terahertz region (THz) are not totally studied and defined yet (Singh et al., 2020) showing great potential for uses in products of biological origin.

THz time-domain spectroscopy (THz-TDS), both in transmission and reflection mode (Wang et al., 2020b), is one of the most applied techniques for the study of the structure and interactions of chemical materials (Zhong et al., 2021); and, according to the optical parameters reported, it is used to discriminate low density materials in food products (Shin et al., 2018), thus becoming a recommended tool for food inspection and quality control, such as fruit and vegetables (Ren et al., 2019a; Ren et al., 2019b).

THz spectroscopy is emerging as a suitable methodology for investigating the dynamic properties of organic molecules relevant to life sciences, because it can detect low-frequency collective vibrational modes of biomolecules (Wilke, 2017).

Terahertz spectroscopy (THz) has been widely used to study spectroscopic characteristics of a variety of materials (Wang et al., 2019), such as semiconductors, biomolecules, liquids, and pharmaceuticals in the spectral region, ranging from 0.1 to 10 THz (Catapano & Soldovieri, 2019; Jiang et al., 2019; Li et al., 2020).

THz spectrum combined with multivariate classification is the strategy used to estimate compounds in the same sample, adulterant in material of different quality or compounds from different samples. The parameters used in THZ investigations spectra (Liu et al., 2020), in order to find differences and have an adequate classification process, are: absorption coefficient, spectral slope of absorption coefficient, and the area of absorption coefficient.

Although many of the spectroscopic techniques mentioned above have been used in the application of food detection, little attention has been paid to the use of Terahertz spectroscopy (THz), which is in a relatively unexplored range of the electromagnetic spectrum ranging from 0.1 to 10 THz, which lies between the mid-infrared and microwave ranges (Catapano & Soldovieri, 2019).

Received 10 Aug., 2022
Accepted 03 Oct., 2022
[1] *Universidad Privada del Norte, Cajamarca, Perú*
*Corresponding author: jimy.oblitas@upn.edu.pe

The composition of cocoa beans is directly influenced by genetic variability, geographical origin and processing. Therefore, chemical and biochemical characteristics and their relationship to external parameters are key characteristics for quality control and technological aspects (Oliveira et al., 2021), Currently, there are studies using near-infrared spectroscopy (NIRS) in the cocoa and chocolate industry (Barbin et al., 2018), showing that it can detect differences but that there are still points for improvement, such as exploring other spectra. Here THz spectroscopy could provide information on time and frequency domains while being insensitive to background thermal radiation (Liu et al., 2018).

Peru is home to six of the main cocoa genetic clusters. Reports such as that by Mejía et al. (2021) show that, in Peruvian varieties of cocoa beans and chocolates, there are significant differences in their chemical and sensory characteristics, making it necessary to investigate techniques such as THz to identify them quickly. This type of spectroscopy will allow improving identification, especially in chocolates, with substitution or addition of ingredients that physicochemically show great differences (Lorenzo et al., 2022), and that need the development of modern techniques that allow their rapid and non-destructive identification.

Looking for the applicability of this technology in the chocolate industry, the objective is to determine the level of differentiation of chocolate bars based on their percentage of cocoa in their composition by using THZ spectroscopy and multivariate analysis.

2 Materials and methods

2.1 Raw material and imaging equipment in the THZ range

The cocoa genotype (*Theobroma cacao* L.) that was used is called "Marañon Native" and comes from the area of Cajamarca -Peru, which was used to make chocolate bars with 50%, 60%, 70%, 80% and 90% of cocoa in their composition. For this process, 10 samples were used for each percentage used. The bars used had dimensions of 10 cm x 10 cm with 0.5 mm of depth. This gave us an image for each sample. In total, 50 images were taken, with 2048 wavelengths, which gave us an average of 600 Mb per square centimeter analyzed.

Terahertz time-domain measurements were obtained by using a Terapulse 4000 spectrometer (Teraview Ltd., Cambridge, UK) in transmission mode. The transmission chamber, the operating scheme are shown in Figure 1. For its operation it was purged with dry nitrogen gas throughout the measurement and the noise was reduced with an average of 10 measurements. Each wave form in the time domain covered a range of 150 ps using a resolution of 0.1 ps. Images were built with equipment scanner. The data acquisition was performed in the TPRJ format and the images were analyzed by using codes internally developed in Matlab v.2019b (Mathworks, Massachusetts, U.S.A.).

2.2 THZ optical parameters

THz optical parameters (refractive index and absorption coefficient) of a sample can be extracted by using the mathematical model proposed by Dhillon et al. (2017). Likewise, to obtain the amplitude and wave phase that is in time domain, a Fourier transform is used to pass it to frequency domain according to the Equation 1 and Equation 2 in Figure 2.

$$E(t) -> FFT -> \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} E(t) e^{-iwt}\, dt = E(w) \tag{1}$$

$$H(w) = \frac{E_S(w)}{E_{ref}(w)} = \frac{4n_0 n_S(w)}{\left[n_S(w) + n_0\right]^2} \cdot \exp\left\{-j\left[n_s(w) - n_0\right]\frac{lw}{c}\right\} \cdot \exp\left[-\frac{lwk_S(w)}{c}\right] \tag{2}$$

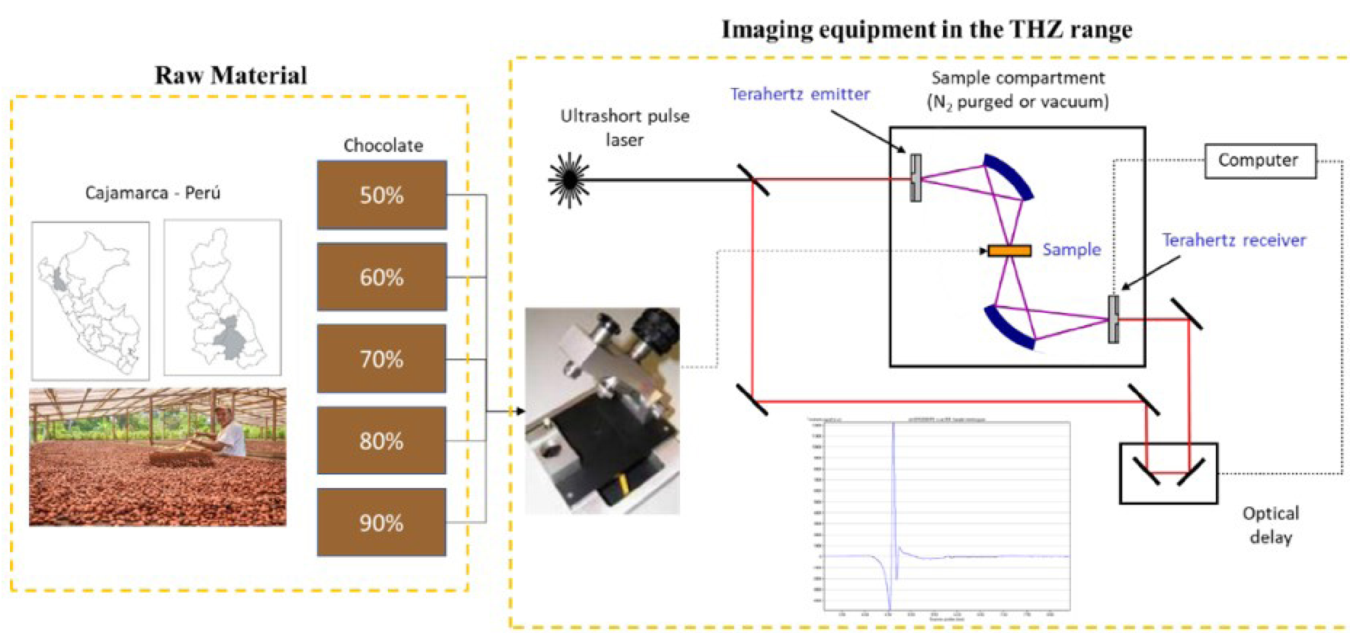


Figure 1. Terahertz pulsed spectroscopy operation schematic.

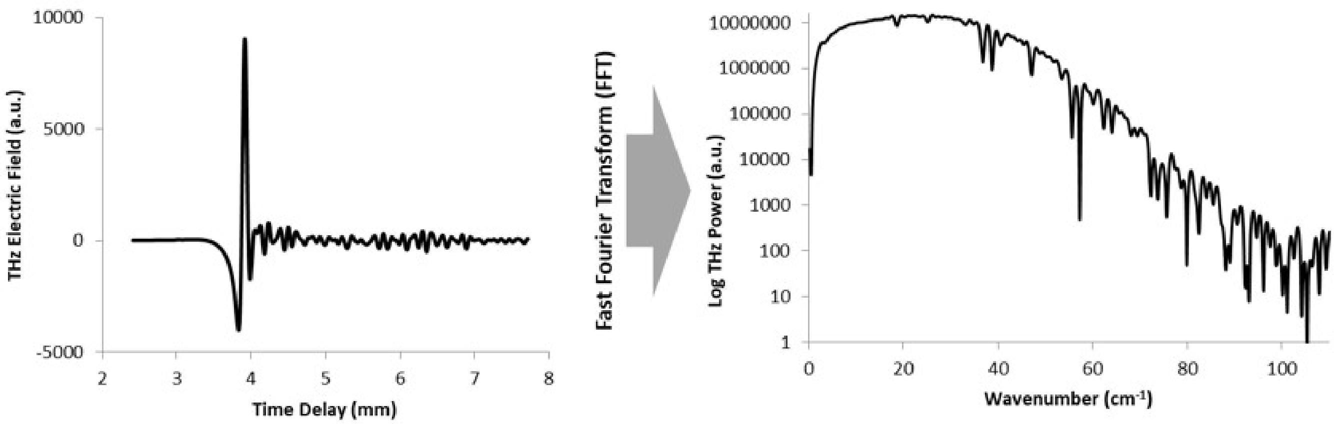


Figure 2. Terahertz data processing.

After Fourier transform, the spectral signal in time domain obtained by the experiment was used to obtain the spectral function in frequency domain, that is E(w). Eref (w) and Es (w) are signals of the reference and the sample, respectively, in frequency domain, and H(w) is the transmission function, which is given by Eq. In Ecs. (1), (2), n_s(w) is the refractive index, Ks(w) is the extinction coefficient, w is the frequency, c is the speed of light, l is the thickness of the sample, and n_0 is the refractive index of nitrogen.

2.3 Multivariate analysis

Multivariate methods can highlight chemical differences among samples and reduce variation due to physical effects. Methods using the Principal Component Analysis (PCA) have shown to be effective for classification, so that they will be applied for this study. PCA was used to obtain a general description of the data, by extracting the main information from the THz spectra of chocolate samples, reducing the number of variables and expressing the total variation in the data set in only a few PCs, with the aim that each spectrum will have its own unique data set, and, therefore, its spectrum can be represented by its PCA scores in factorial space.

In addition, because PCA are orthogonal representations, we can represent it graphically and observe the relationships between them, and provide very important information on the potential capacity to differentiate chocolate samples based on their percentage of cocoa. To identify the best classifier, 7 classification algorithms were used: decision trees, linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), support vector machines (SVM), closer neighbor classifiers, set classifiers, and naive Bayesian classifiers. Each of these algorithms was combined with its level of interpretability and flexibility, obtaining 24 models that are used in this research. The selection of discriminatory variables was carried out by using a feature selection technique based on the staggered decoupling of variables.

Cross-validation was used by randomly dividing the original THZ spectra data set into a training set and a test set, measuring the mean cross-validation error as a performance indicator. For the other parameters, a heuristic procedure was used to select the scale value based on Kernel function to calculate the best classifier. The best model shall be determined based on its accuracy.

3 Results and discussion

3.1 Spectra analysis

A spectral image is taken in time domain, ranging from 0 to 10 THZ, from which the values of each pixel analyzed in chocolate samples were collected. Figure 3 shows the average of values obtained.

THZ pulses shown in Figure 3 demonstrate that the reference pulse is in advanced with respect to the signals of each of the chocolate samples. Besides, the amplitude decreases due to the phenomena of refraction and absorption of the samples. It is also interesting to note that the amplitude of THz time-domain signal transmission is proportional to the cocoa content in the samples. This may also be due to moisture interactions the same ones could have.

Figure 4 shows the transmitted THz amplitude spectra obtained by the Fourier transform of time domain signals.

Figure 4 shows the Fourier transforms ranging from 0 to 2 THZ. It is observed in this field that amplitudes decrease as cocoa content in the samples decrease, due to the different substantial absorption of the samples, being the largest difference in the range from 0 to 1.4 THZ. It is also important to note that the peaks shown by the reference are replicated in chocolate samples. These are due to the ambient humidity conditions where the sample was taken.

The actual dielectric function of chocolate samples was also obtained, observing that only in the range from 0.2 to 1.6 a differentiation of samples is achieved (Figure 5). This is important since this property shows a direct relation with the moisture content of the samples and the possible structural changes that chocolate may have in its shelf life.

High water content in tissues and biological liquids determine the nature of their dielectric response in the Terahertz frequency range (THz) (Smolyanskaya et al., 2018). This is the case for

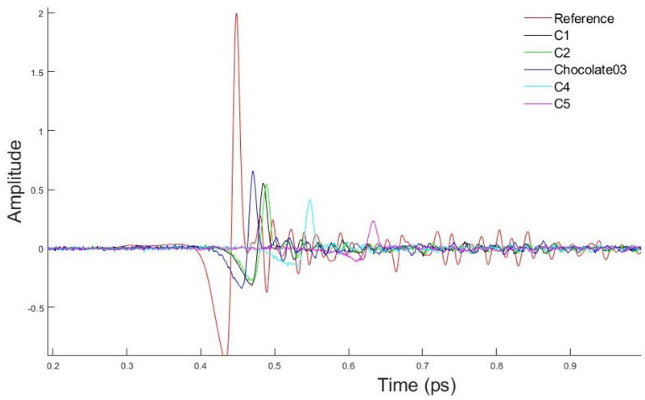


Figure 3. Diagram of THz frequency-domain spectroscopy.

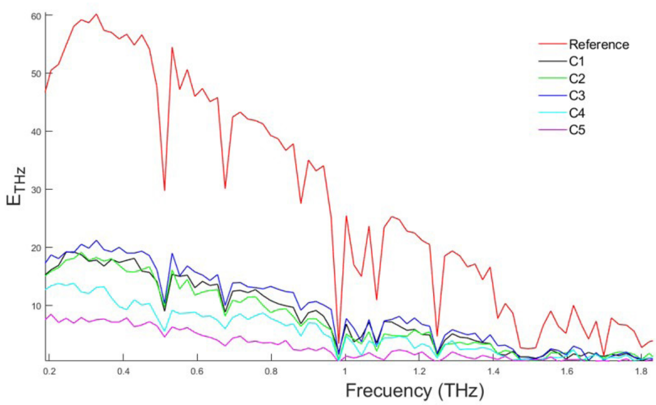


Figure 4. Corresponding THz pulses in frequency domain.

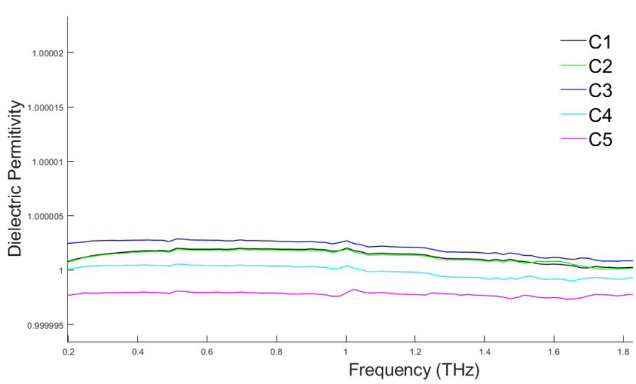


Figure 5. Actual dielectric function.

many food products ranging from dry products to products with high water content. This part of products' composition has free or bound water. These two types make valuable but different contributions to the THz response of biological objects (Cherkasova et al., 2020).

Absorption coefficients based on THZ frequency are also shown, although it was analyzed up to 10 THZ. Only in the range shown, it is observed a difference between chocolate samples. It is also observed that there is a slight increase in absorbance coefficients as the frequency increases (Wang et al., 2020a). This is given in all samples and that of 90% of cocoa has the highest rates.

3.2 Multivariate analysis

In products, such as chocolate, it is difficult to discriminate its cocoa content, and, since this quality characteristic is related to its cost, it is necessary to consider alternatives that can adequately discriminate it. At this point, the analysis of curves in frequency domain is shown as an alternative, since multivariant methods could detect significant differences that allow us to generate an appropriate discrimination model between chocolates with different cocoa content.

Since one of the parameters with the greatest understanding is absorbance, it will be used to perform the multivariate analysis. The respective curves showing a difference were already shown in Figure 6, with the exception of some points, such as the frequency of 1.6 THZ showing an overlap. The values between 0.1 and 2 THZ shall be used for multivariate analysis.

All images were pretreated with a linear filter to reduce image noise, as recommended by Guo et al. (2018). The THz spectral imaging data set of the sample is based on specific parameters, such as time interval, amplitude or THz wave phase, and, then, makes the refractive index, spatial density distribution, thickness distribution, and contour of the sample.

After data pretreatment and statistical adjustment, 24 classification models were tested by using linear and non-linear models, taking accuracy percentage as an indicator. All models used a 15-fold cross-validation. The summary of the best models is shown in Figure 7.

The best model was the optimized Fine Gaussian SVM model that obtained a Precision of 93%, with a Kernel Scale of 1 and cubic function type, and a Multiclass Method One vs One, optimized with a Bayesian function of 30 interactions. This type of model has been reported many times in the research on the use of Machine Learning for image recognition (Loussaief & Abdelkrim, 2018).

Figure 8a shows the score graph of the main components of absorption spectra of chocolate samples, which have most of the spectral variations with PC1 (63.8%) and PC2 (36.2%). As noted in this figure, samples are clearly divided into 5 groups according to their cocoa content in composition. PC1 has almost all the capacity for discrimination by confirming that Terahertz spectra have sufficient information to classify different products according to their composition.

A new sample was tested by using the generated classification model. Results are shown in the confusion matrix of Figure 8b, where it can be observed that the classification level of chocolate samples is satisfactory, generating few false positive errors.

In order to test discrimination efficiency in ranges higher than 2 THZ, it was analyzed following the same procedure, reaching unsatisfactory values in the capacity to classify chocolates for their cocoa content. Studies on classification of rice (Xu et al., 2015) and soybean (Liu et al., 2016) by using THZ spectroscopy techniques have been reported with great success in classification.

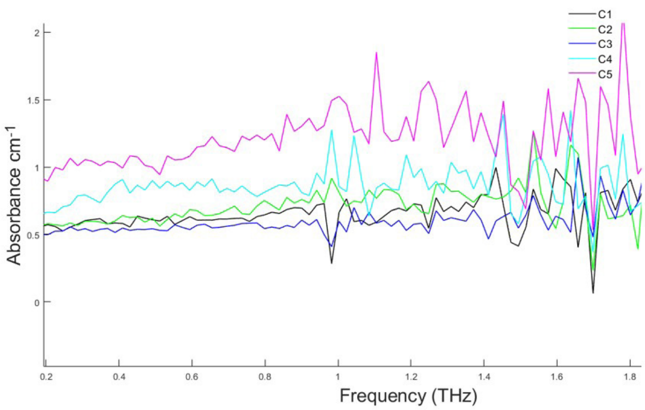


Figure 6. Absorbance of the samples.

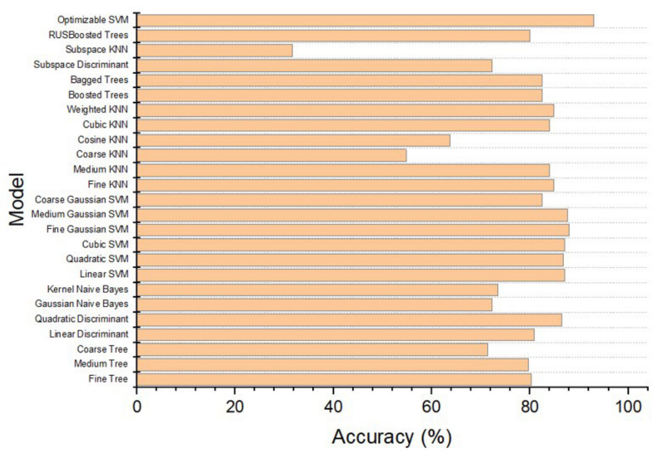
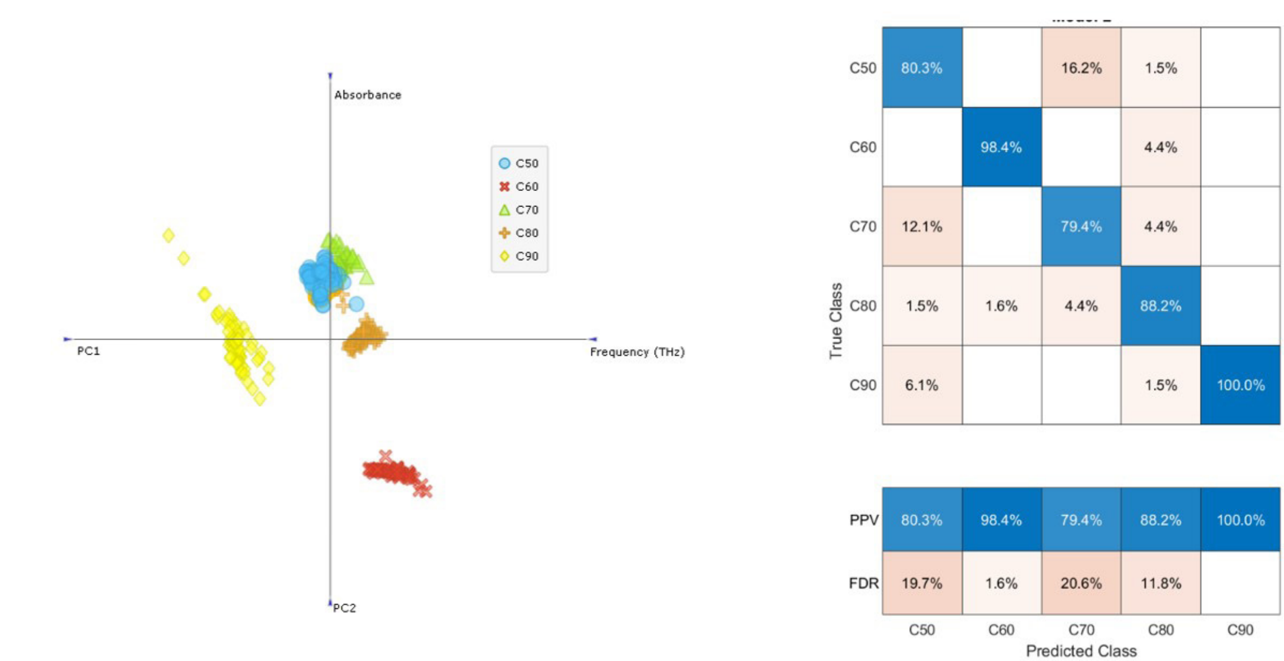


Figure 7. Models with the best accuracy.

According to Mahanta et al. (2019), by only finding differences that can generate an adequate classification in the range lower than 2 THZ, it can be said that this might be due to "fast relaxation" processes as well as to hidden water bending mode at 1.5 THz. These theories, while still under discussion, are linked to bound and free water content in biological matrices, such as chocolate; and, since this product has very low water content, it can be foreseen that the detection of differences at this level of ranges is related to bound water. This appreciation shows that the THZ spectrum can be very useful in the concept of relating this type of water to food composition.

The chocolate industry is most likely one of those which could suffer problems of non-compliance with the labels regarding cocoa content, so that analyzing the new spectra to see its potential for classification becomes an issue of importance. Similar studies in chocolate bars contaminated with foreign materials showed great efficiency by using the THZ range spectra (Weiller et al., 2018). These results show a great capacity to detect and discriminate different types of materials according to their composition, especially, since, within food industry, the analysis of a component is always an expensive, slow and destructive technique. Even then, other spectroscopic techniques such as mass spectroscopy are also being used (Lorenzo & Pico, 2017), so the importance of evaluating little-studied ranges in food, as given in Time-Domain Spectroscopy, becomes necessary, especially appreciating the importance of non-destructive inspection for food industry, meeting the needs of modern and rapid techniques for international food trade.

Another great advantage of using THZ spectroscopy is that it has the capacity to penetrate the test substance (Wang et al., 2019). This is particularly important in food-based materials,



(a) (b)

Figure 8. (a) PCA score graph; (b) confusion matrix.

such as chocolate, as these products often form different colors on their surface as a reaction with air, which can generate a wrong reading by more direct techniques, such as NIR spectroscopy or images of multiple spectra.

In this research, it is tested the suitability of this technology for classifying chocolate bars according to the cocoa content. Although the results are very encouraging, THz-TDS still remains a technology that involves high implementation costs. However, the recent and rapid development of THz systems for agro-industrial research opens up real possibilities for these costs to be significantly reduced in the coming years.

4 Conclusion

Finally, it can be concluded that THz-TDS spectroscopy, together with multivariate classification techniques, such as PCA, have the capacity to classify chocolate bars according to their percentage of cocoa. This is a promising technique, because, besides the high classification efficiency, samples do not need any important physical pretreatment for analysis. These characteristics give this technique great potential for its application in food quality control processes.

References

Balthazar, C. F., Guimarães, J. T., Rocha, R. S., Pimentel, T. C., Cucinelli, R. P. No., Tavares, M. I. B., Graça, J. S., Alves, E. G. Fo., Freitas, M. Q., Esmerino, E. A., Granato, D., Rodrigues, S., Raices, R. S. L., Silva, M. C., Sant'Ana, A. S., & Cruz, A. G. (2021). Nuclear magnetic resonance as an analytical tool for monitoring the quality and authenticity of dairy foods. *Trends in Food Science & Technology*, 108, 84-91. http://dx.doi.org/10.1016/j.tifs.2020.12.011.

Barbin, D. F., Maciel, L. F., Bazoni, C. H. V., Ribeiro, M. S., Carvalho, R. D. S., Bispo, E. S., Miranda, M. P. S., & Hirooka, E. Y. (2018). Classification and compositional characterization of different varieties of cocoa beans by near infrared spectroscopy and multivariate statistical analyses. *Journal of Food Science and Technology*, 55(7), 2457-2466. http://dx.doi.org/10.1007/s13197-018-3163-5. PMid:30042561.

Castro, W., Oblitas, J., Chuquizuta, T., & Avila-George, H. (2017). Application of image analysis to optimization of the bread-making process based on the acceptability of the crust color. *Journal of Cereal Science*, 74, 194-199. http://dx.doi.org/10.1016/j.jcs.2017.02.002.

Catapano, I., & Soldovieri, F. (2019). THz imaging and data processing: state of the art and perspective. In R. Persico, S. Piro & N. Linford (Eds.), *Innovation in near-surface geophysics: instrumentation, application, and data processing methods* (pp. 399-417). Amsterdam: Elsevier. http://dx.doi.org/10.1016/B978-0-12-812429-1.00011-8.

Cherkasova, O. P., Nazarov, M. M., Konnikova, M., & Shkurinov, A. P. (2020). THz Spectroscopy of bound water in glucose: direct measurements from crystalline to dissolved state. *Journal of Infrared, Millimeter and Terahertz Waves*, 41(9), 1057-1068. http://dx.doi.org/10.1007/s10762-020-00684-4.

De-la-Torre, M., Avila-George, H., Oblitas, J., & Castro, W. (2020). Selection and fusion of color channels for ripeness classification of cape gooseberry fruits. In J. Mejia, M. Muñoz, Á. Rocha & J. A. Calvo-Manzano (Eds.), *Trends and applications in software engineering* (pp. 219-233, Advances in Intelligent Systems and Computing, Vol. 1071). Cham: Scopus. http://dx.doi.org/10.1007/978-3-030-33547-2_17.

Dhillon, S. S., Vitiello, M. S., Linfield, E. H., Davies, A. G., Hoffmann, M. C., Booske, J., Paoloni, C., Gensch, M., Weightman, P., Williams, G. P., Castro-Camus, E., Cumming, D. R. S., Simoens, F., Escorcia-Carranza, I., Grant, J., Lucyszyn, S., Kuwata-Gonokami, M., Konishi, K., Koch, M., Schmuttenmaer, C. A., Cocker, T. L., Huber, R., Markelz, A. G., Taylor, Z. D., Wallace, V. P., Zeitler, J. A., Sibik, J., Korter, T. M., Ellison, B., Rea, S., Goldsmith, P., Cooper, K. B., Appleby, R., Pardo, D., Huggard, P. G., Krozer, V., Shams, H., Fice, M., Renaud, C., Seeds, A., Stöhr, A., Naftaly, M., Ridler, N., Clarke, R., Cunningham, J. E., & Johnston, M. B. (2017). The 2017 terahertz science and technology roadmap. *Journal of Physics. D, Applied Physics*, 50(4), 043001. http://dx.doi.org/10.1088/1361-6463/50/4/043001.

Guo, L., Chen, L., Chen, C. L. P., & Zhou, J. (2018). Integrating guided filter into fuzzy clustering for noisy image segmentation. *Digital Signal Processing*, 83, 235-248. http://dx.doi.org/10.1016/j.dsp.2018.08.022.

Jiang, Y., Ge, H., & Zhang, Y. (2019). Detection of foreign bodies in grain with terahertz reflection imaging. *Optik*, 181, 1130-1138. http://dx.doi.org/10.1016/j.ijleo.2018.12.066.

Li, C., Li, B., & Ye, D. (2020). Analysis and identification of rice adulteration using Terahertz spectroscopy and pattern recognition algorithms. *IEEE Access*, 8, 26839-26850. http://dx.doi.org/10.1109/ACCESS.2020.2970868.

Liu, W., Liu, C., Chen, F., Yang, J., & Zheng, L. (2016). Discrimination of transgenic soybean seeds by terahertz spectroscopy. *Scientific Reports*, 6(1), 35799. http://dx.doi.org/10.1038/srep35799. PMid:27782205.

Liu, W., Liu, C., Yu, J., Zhang, Y., Li, J., Chen, Y., & Zheng, L. (2018). Discrimination of geographical origin of extra virgin olive oils using terahertz spectroscopy combined with chemometrics. *Food Chemistry*, 251, 86-92. http://dx.doi.org/10.1016/j.foodchem.2018.01.081. PMid:29426428.

Liu, W., Zhang, Y., Li, M., Han, D., & Liu, W. (2020). Determination of invert syrup adulterated in acacia honey by terahertz spectroscopy with different spectral features. *Journal of the Science of Food and Agriculture*, 100(5), 1913-1921. http://dx.doi.org/10.1002/jsfa.10202. PMid:31846080.

Lorenzo, M., & Pico, Y. (2017). Gas chromatography and mass spectroscopy techniques for the detection of chemical contaminants and residues in foods. In D. Schrenk & A. Cartus (Eds.), *Chemical contaminants and residues in food* (2nd ed., pp. 15-50). Oxford: Woodhead Publishing. http://dx.doi.org/10.1016/B978-0-08-100674-0.00002-3.

Lorenzo, N. D., Santos, O. V., & Lannes, S. C. S. (2022). Structure and nutrition of dark chocolate with pequi mesocarp (*Caryocar villosum* (Alb.) Pers.). *Food Science and Technology*, 42, e88021. http://dx.doi.org/10.1590/fst.88021.

Loussaief, S., & Abdelkrim, A. (2018). Machine learning framework for image classification. *Advances in Science, Technology and Engineering Systems*, 3(1), 1-10. http://dx.doi.org/10.25046/aj030101.

Mahanta, D.D., Islam, S. I., Choudhury, S., Das, D. K., Mitra, R. K., & Barman, A. (2019). Contrasting hydration dynamics in DME and DMSO aqueous solutions: a combined optical pump-probe and GHz-THz dielectric relaxation investigation. *Journal of Molecular Liquids*, 290, 111194. http://dx.doi.org/10.1016/j.molliq.2019.111194.

Mejía, A., Meza, G., Espichán, F., Mogrovejo, J., & Rojas, R. (2021). Chemical and sensory profiles of Peruvian native cocoas and chocolates from the Bagua and Quillabamba regions. *Food Science and Technology* 41(Suppl. 2), 576-582. http://dx.doi.org/10.1590/fst.08020.

Oliveira, A. P. F., Milani, R. F., Efraim, P., Morgano, M. A., & Tfouni, S. A. V. (2021). Cd and Pb in cocoa beans: occurrence and effects of chocolate processing. *Food Control*, 119, 107455. http://dx.doi.org/10.1016/j.foodcont.2020.107455.

Ren, A., Zahid, A., Imran, M. A., Alomainy, A., & Abbasi, Q. H. (2019a). Introducing a novel technique of detecting fruits contaminations using Terahertz sensing. In *2019 International Workshop on Antenna Technology (iWAT)* (pp. 58-61). Manhattan: IEEE. http://dx.doi.org/10.1109/IWAT.2019.8730602.

Ren, A., Zahid, A., Imran, M. A., Alomainy, A., Fan, D., & Abbasi, Q. H. (2019b). Terahertz sensing for fruit spoilage monitoring. In F. Zheng (Ed.), *2019 Second International Workshop on Mobile Terahertz Systems (IWMTS)* (pp. 1-4). Manhattan: IEEE. https://doi.org/10.1109/IWMTS.2019.8823735.

Shin, H. J., Choi, S.-W., & Ok, G. (2018). Qualitative identification of food materials by complex refractive index mapping in the terahertz range. *Food Chemistry*, 245, 282-288. http://dx.doi.org/10.1016/j.foodchem.2017.10.056. PMid:29287372.

Singh, A. K., Pérez-López, A. V., Simpson, J., & Castro-Camus, E. (2020). Three-dimensional water mapping of succulent *Agave victoriae-reginae* leaves by terahertz imaging. *Scientific Reports*, 10(1), 1404. http://dx.doi.org/10.1038/s41598-020-58277-z. PMid:31996722.

Smolyanskaya, O. A., Chernomyrdin, N. V., Konovko, A. A., Zaytsev, K. I., Ozheredov, I. A., Cherkasova, O. P., Nazarov, M. M., Guillet, J.-P., Kozlov, S. A., Kistenev, Y. V., Coutaz, J.-L., Mounaix, P., Vaks, V. L., Son, J.-H., Cheon, H., Wallace, V. P., Feldman, Y., Popov, I., Yaroslavsky, A. N., Shkurinov, A. P., & Tuchin, V. V. (2018). Terahertz biophotonics as a tool for studies of dielectric and spectral properties of biological tissues and liquids. *Progress in Quantum Electronics*, 62, 1-77. http://dx.doi.org/10.1016/j.pquantelec.2018.10.001.

Wang, C., Zhou, R., Huang, Y., Xie, L., & Ying, Y. (2019). Terahertz spectroscopic imaging with discriminant analysis for detecting foreign materials among sausages. *Food Control*, 97, 100-104. http://dx.doi.org/10.1016/j.foodcont.2018.10.024.

Wang, J., Wu, X., Zheng, J., & Wu, B. (2022). Rapid identification of green tea varieties based on FT-NIR spectroscopy and LDA/QR. *Food Science and Technology*, 42, e73022. http://dx.doi.org/10.1590/fst.73022.

Wang, K., Sun, D.-W., & Pu, H. (2017). Emerging non-destructive terahertz spectroscopic imaging technique: principle and applications in the agri-food industry. *Trends in Food Science & Technology*, 67, 93-105. http://dx.doi.org/10.1016/j.tifs.2017.06.001.

Wang, Q., Hameed, S., Xie, L., & Ying, Y. (2020a). Non-destructive quality control detection of endogenous contaminations in walnuts using terahertz spectroscopic imaging. *Journal of Food Measurement and Characterization*, 14(5), 2453-2460. http://dx.doi.org/10.1007/s11694-020-00493-2.

Wang, Y., Zhao, Z., Qin, J., Liu, H., Liu, A., & Xu, M. (2020b). Rapid in situ analysis of l-histidine and α-lactose in dietary supplements by fingerprint peaks using terahertz frequency-domain spectroscopy. *Talanta*, 208, 120469. http://dx.doi.org/10.1016/j.talanta.2019.120469. PMid:31816746.

Weiller, S., Tanabe, T., & Oyama, Y. (2018). Terahertz non-contact monitoring of cocoa butter in chocolate. *World Journal of Engineering and Technology*, 6(2), 268-274. http://dx.doi.org/10.4236/wjet.2018.62015.

Wilke, I. (2017). Terahertz spectroscopy applications. In P. Worsfold, C. Poole, A. Townshend & M. Miró (Eds.), *Encyclopedia of analytical science* (3rd ed., pp. 1-5). Amsterdam: Academic Press.

Xu, W., Xie, L., Ye, Z., Gao, W., Yao, Y., Chen, M., Qin, J., & Ying, Y. (2015). Discrimination of transgenic rice containing the Cry1Ab protein using Terahertz spectroscopy and chemometrics. *Scientific Reports*, 5(1), 11115. http://dx.doi.org/10.1038/srep11115. PMid:26154950.

Xu, Y., Zhong, P., Jiang, A., Shen, X., Li, X., Xu, Z., Shen, Y., Sun, Y., & Lei, H. (2020). Raman spectroscopy coupled with chemometrics for food authentication: a review. *Trends in Analytical Chemistry*, 131, 116017. http://dx.doi.org/10.1016/j.trac.2020.116017.

Zhong, J., Mori, T., Kashiwagi, T., Yamashiro, M., Kusunose, S., Mimami, H., Tsujimoto, M., Tanaka, T., Kawashima, H., Nakagawa, S., Ito, J., Kijima, M., Iji, M., Watanabe, M. M., & Kadowaki, K. (2021). Characteristic terahertz absorption spectra of paramylon and paramylon-ester compounds. *Spectrochimica Acta. Part A: Molecular and Biomolecular Spectroscopy*, 244, 118828. http://dx.doi.org/10.1016/j.saa.2020.118828. PMid:32882654.