

A method for non-destructive determination of cocoa bean fermentation levels based on Terahertz hyperspectral imaging

Dinh T. NGUYEN^{1,4}, Audrey PISSARD², Juan Antonio FERNANDEZ PIERNA², Hervé ROGEZ³,
Jesus SOUZA³, Fabian DORTU¹, Saurav Goel^{4,5,6}, Yves HERNANDEZ¹ and Vincent BAETEN²*

¹*Applied Photonics Department, Multitel A.S.B.L, Rue Pierre et Marie Curie 2, Mons 7000, Belgium*

²*Quality and authentication Unit, Knowledge and valorization of agricultural products Department, Walloon
Agricultural Research Centre, 24 chaussée de Namur, 5030 Gembloux, Belgium*

³*CVACBA (Center for Valorization of Amazonian Bioactive Compounds), UFPA (Federal University of Pará),
Avenida Perimetral, 01, 66075-150, Guamá, Belém, Pará, Brazil*

⁴*School of Engineering, London South Bank University, London, SE10AA, United Kingdom*

⁵*Indian Institute of Technology Guwahati, Guwahati, 781039, India*

⁶*University of Petroleum and Energy Studies, Dehradun, 248007, India*

* Corresponding author: nguyed13@lsbu.ac.uk

ABSTRACT

Fermentation of cocoa is a key process to obtain aromatic chocolate products from raw cocoa beans. Hitherto, the levels of fermentation in cocoa are determined using destructive techniques, for example by a cut-test to manually observe the colour inside the beans, or by quantifying ammonia nitrogen (NH₃) in the cocoa powder. In this paper, we present the use of Terahertz hyperspectral imaging as a new way to non-destructively analyse and detect fermented cocoa beans. The study analysed two sets of twenty-two cocoa bean samples with different levels of fermentation from two producers in Brazil. A correlation between fermentation conditions and the outcome results of their THz measurements was observed.

Keywords: *Quality control of cocoa beans, THz imaging, Principal component analysis.*

26 **1. Introduction**

27 Chocolate flavour has been used in the food sector since the discovery of fermentation of
28 cocoa beans. An increase by over 3% demand of cocoa use across the world since 2008 and
29 the production increase from 2.85 million ton in 2000-2001 to 4.84 million tons speaks for
30 the rising demand in this food sector (Beg et al., 2017).

31 Choosing a premier quality cocoa is an important attribute in the chocolate production. The
32 characteristic of final chocolate products is linked to the fermentation process of cocoa beans.
33 The fermentation quality is dependent on fermentation time (days), temperature or drying
34 method and geographical origin, which governs and influence the quality and taste of the
35 chocolate. Hitherto, it has continued to remain a challenge for the chocolate manufacturers to
36 identify and rapidly screen good quality cocoa.

37 Nowadays, evaluating commercial grade cocoa beans requires to perform a destructive
38 cutting test of the cocoa beans to count the proportion of purple and brown beans on
39 representative dried samples (Beans, 2015; Santos et al., 2019). However, the cut test is not
40 engineering accurate method and labour oriented as it needs manual counting. For example,
41 (Ilangantileke et al., 1991) presented results of the cut-test score and sensory evaluation
42 methods and it highlighted the variability in results and inadequacies in assessment of the
43 cocoa bean quality.

44 Another popular method that is commonly used for enhanced quality control of cocoa beans
45 is by quantification of ammonia nitrogen (NH_3) content in the cocoa powder generated during
46 the fermentation due to the degradation of proteins. It is generally quantified by the
47 traditional Conway method (Conway and Flood, 1936). The Conway method measures the
48 absorption of NH_3 volatile matter (VoC) in the air by some reagents (e.g., an acid) which
49 leads to change in the pH value. An alternative to the Conway method is by determining

50 different NH₃ levels in the cocoa powder by using reflectance NIR spectroscopy (Hue et al.,
51 2014). All these methods are destructive and time consuming and can take even upto 48h.
52 NIR and Fourier Transform NIR (FT-NIR) techniques are quite popular even to estimate the
53 quality of cocoa for example, variety, composition and fraud detection (Barbin et al., 2018;
54 Quelal-Vásconez et al., 2019; Quelal-Vásconez et al., 2018; Sunoj et al., 2016; Teye et al.,
55 2015). However, these methods cannot be considered as ideal as they require the process of
56 powdering the beans, i.e., destroying the cocoa beans before performing the measurements.
57 More recent alternatives include the use of NIR Hyperspectral Imaging that combines NIR
58 spectroscopy and spatial information of the samples. NIR hyperspectral imaging can be
59 directly applied on single cocoa bean for non-destructive prediction of fermentation index,
60 polyphenol content and antioxidant activity (Caporaso et al., 2018) or for the authentication
61 of cocoa bean hybrids (Cruz-Tirado et al., 2020), among others.
62 Imaging can be also combined with other spectroscopic techniques such as Terahertz (THz).
63 Both THz spectroscopy and imaging are well known non-destructive testing and analysis
64 techniques used in biological (Assefzadeh et al., 2016) and agro-food products, such as the
65 detection of cocoa butter in chocolate (Weiller et al., 2018) and food spoilage evaluation
66 (Hindle et al., 2018). The advantage of THz radiation is that it can reach deeper into the
67 subsurface of non-polar and opaque materials compared to near and mid infrared (IR) ranges.
68 In THz time-domain spectroscopy (THz-TDS) measurements, the electric field amplitude is
69 sampled as a function of time, providing both magnitude and phase spectra. By scanning the
70 sample under the beam, 2D images can be recorded. Valuable information such as refractive
71 index, extinction coefficient, etc. can be extracted from a single THz-TDS measurement in a
72 rapid timeframe. Moreover, recent studies show that NH₃ has fingerprint absorption
73 characteristics in THz range (Assefzadeh et al., 2016) , which opens up an opportunity in this
74 work for quantifying fermentation of cocoa beans using THz technology.

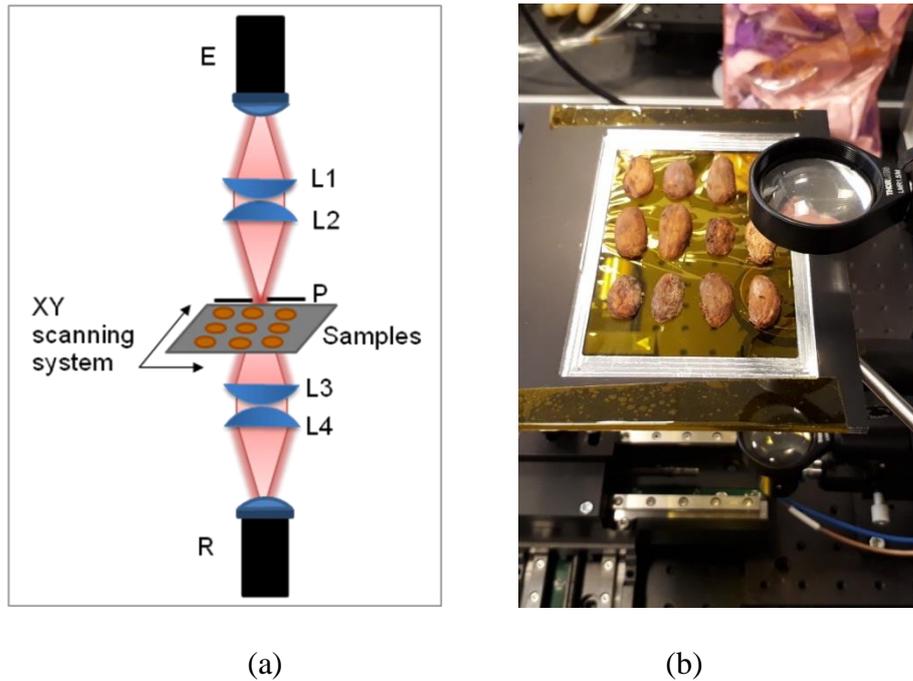
75 In this paper, we present an approach for using THz time-domain spectroscopy as a novel
76 way for non-destructive evaluation of cocoa bean fermentation levels. For this study, two sets
77 of twenty-two cocoa beans/each processed through different fermentation times from two
78 producers in Brazil were used. The first set of twenty-two beans were used to perform a
79 preliminary study using the THz technology and the second set was used for validation to
80 certify the reproducibility of the results obtained during the first sets of study. For this, THz
81 hyperspectral images of different sets of samples were recorded and analysed using simple
82 statistical models as well as the multivariate analysis e.g., Principal Component Analysis
83 (PCA) method. A correlation between fermentation conditions and the outcome results of
84 their THz measurements was observed which is presented in the results section.

85

86 **2. Materials and method**

87 *2.1. Instrumental setup*

88 This study made use of a THz-TDS system to record hyperspectral images of the samples
89 in transmission mode. The experimental setup is schematically presented in *Fig. 1(a)*. A
90 mode-locked femtosecond laser (with 100 fs pulse duration, 50 Hz repetition rate, 1550 nm
91 wavelength, and 100 mW average power) was used to pump the photoconductive antenna
92 emitter (*E*) and gate the detector antenna (*R*). A delay line changes the optical path difference
93 between THz and the probe beams, combining coherently at the detector to sample the
94 impulse response signal.



95

96

97 **Figure 1.** (a) Schematic diagram of the TDS system used for cocoa bean measurements in the transmission
 98 mode. *E* and *R* are THz photoconductive antennas for emitter and receiver respectively. *L1*, *L2*, *L3*, *L4* are
 99 identical plano-convex lenses. *P* is a pinhole. (b) Photograph of the cocoa semi-bean samples mounted on the
 100 measurement platform.

101 Each measurement was performed at room conditions and in the transmission mode,
 102 through a 4-f optical setup using four plano-convex lenses (*L1-L4*) of the same focal length
 103 and made of Polymethylpentene material (TPX). Hyperspectral imaging data was obtained by
 104 scanning the samples point-by-point. Further details of this instrument setup can be assessed
 105 from an alternative source (Nguyen et al., 2018).

106 A possible mechanism influencing the THz spectroscopic results can be levels of humidity
 107 present in the room as THz waves are extremely sensitive to the presence of water. To avoid
 108 the effect of humidity on the THz peak-to-peak intensity, a short optical pathlength was
 109 deliberately used by choosing appropriate optics which may not be a highly optimal
 110 configuration.

111 2.2. Sample collection

112 The study used two sets of twenty-two cocoa beans per set processed through different

113 fermentation times (0, 2 and 4 or 7 days). The first set (SET1) of samples was measured in
 114 December 2019 and the second set (SET2 or validation set) of samples was measured in June
 115 2020. SET1 was used to perform a preliminary study of THz technology for this kind of data,
 116 and SET2 was used as validation to certify the reproducibility of the results obtained with
 117 SET1.

118 The cocoa bean samples were previously cut in half leading to two ‘semi-beans’. This step
 119 was carried out so that we have one half to perform THz measurements, whilst the other half
 120 was kept for further analysis, but the technique also has the potential for its use in a non-
 121 destructive manner, e.g., one can always measure the full samples non-destructively by THz
 122 measurements. As an exploratory study, the sample details used in this study as SET1 and
 123 SET2 are listed in Table 1 and 2 respectively. For each type of sample used, two semi-beans
 124 from same fermentation conditions were available for the THz measurements. For every pair
 125 of samples, information about fermentation time, height of the fermented layer, as well as the
 126 method of drying (mode and time) are shown. Therefore, the THz measurements were
 127 realized on ‘semi-bean’ samples, which were placed on the transparent thin plastic film for
 128 2D scanning, as shown in Figure 1 (b).

129 **Table 1.** Sample list of SET1.

Sample No.	Fermentation time (days)	Height of fermented layer (cm)	Drying (mode)	Time of drying (days)
1-2	4	20	Not Dried	0
3-4	4	50	Not Dried	0
5-6	7	60	Not Dried	0
7-8	7	60	Shadow	4
9-10	7	60	Sun	3
11-12	7	60	Not Dried	0
13-14	7	60	Sun	3

15-16	0	0	Shadow	3
17-18	0	0	Sun	3
19-20	0	60	Not Dried	0
21-22	0	60	Not Dried	0

130

131 **Table 2.** *Sample list of SET2*

Sample No.	Fermentation time (days)	Height of fermented layer (cm)	Drying (mode)	Time of drying (days)
1-2	0	0	Shadow	4
3-4	0	0	Sun	1
5-6	0	NaN	Not Dried	0
7-8	4	40	Not Dried	0
9-10	4	30	Not Dried	0
11-12	7	60	Shadow	4
13-14	7	60	Not Dried	0
15-16	7	60	Not Dried	0
17-18	2	40	Not Dried	0
19-20	2	30	Not Dried	0
21-22	2	60	Not Dried	0

132

133 With the current THz-TDS setup using mechanical delay line, the samples were scanned
134 with a spatial scanning step of 0.4 mm, and it took approximately 6 minutes to finish one
135 seed measurement for the first set and the measurement time was successfully reduced to 4
136 minutes/sample for the validation set by optimisation of the scanning zone. Scanning time
137 could further be reduced by fully optimising the scanning step. However, the interest of this
138 study is not to optimise the agility of measurement but to obtain the accuracy of
139 measurements and its applicability by performing the feasibility tests.

140 2.3. Data analysis method

141 Figure 2 (a) depicts a recorded THz-TDS pulse signal at a single pixel point, from which one
 142 can extract the peak-to-peak (*ptp*) magnitude information as follow:

143

$$144 \quad \mathbf{ptp} = \mathbf{max}(E_{THz}) - \mathbf{min}(E_{THz}) \quad (1)$$

145

146 where E_{THz} is the recorded THz electromagnetic field obtained from the THz-TDS system.

147 Full peak-to-peak image was then mapped in 2D showing substantial contrast depending on

148 the sample properties, as shown in Figure 2(b). Another information that was extracted from

149 the measured THz-TDS pulse signal is the delay time difference between a measured pixel

150 point and that of a reference (the transparent thin plastic film without bean sample). The time-

151 delay difference represents a group refractive index, and can be estimated as:

$$152 \quad \Delta t = t_{maxSample} - t_{maxRef} \quad (2)$$

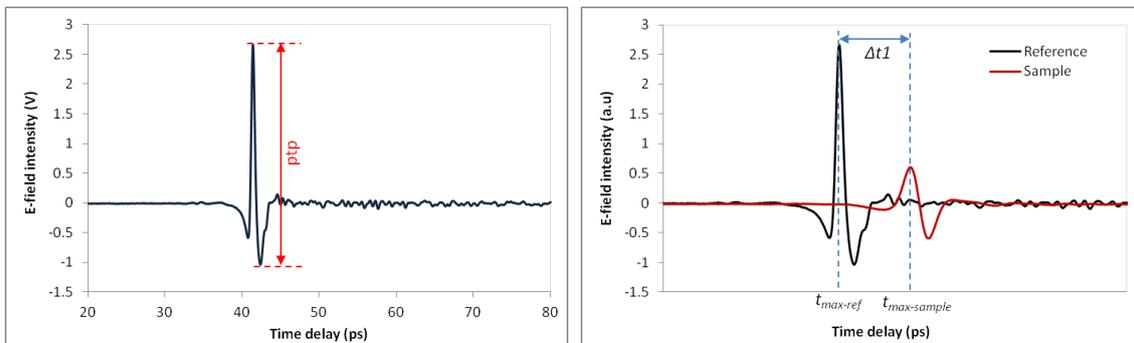
153 where $t_{maxSample}$ is the time-delay (picoseconds) of the maximum measured E-field of the

154 sample whereas t_{maxRef} is that of the reference, and Δt is the difference between these two

155 parameters. Figure 2(c) illustrates how this delay-time difference was calculated.

156 Figure 2(d) shows a map of the signal amplitude as a function of y position along a line

157 through the cocoa bean and the time delay.

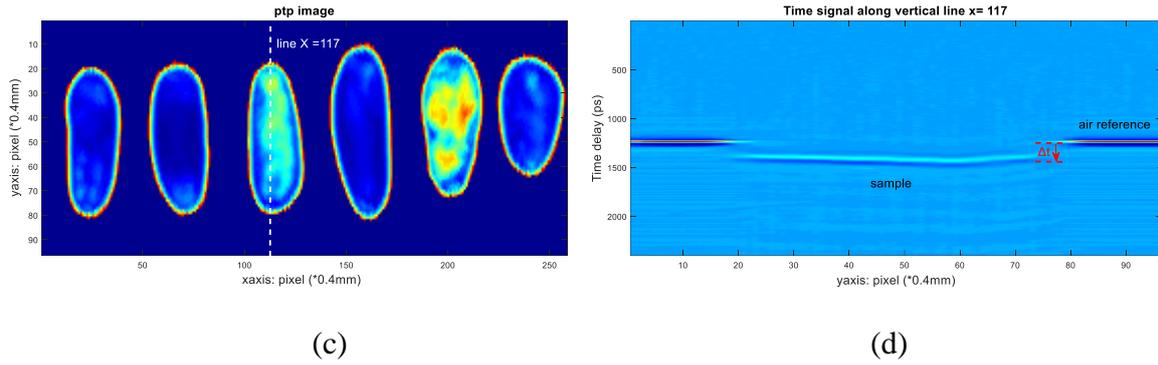


158

(a)

(b)

159



160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

Figure 2. (a, b) illustrations of the definitions of the peak-to-peak amplitude and peak-to-peak image of 6 samples; (c,d) illustrations of definition of time-delay difference between sample and reference and time-delay along the marked vertical line in (b). The colour variation in 2(c) indicates the signal intensity received at the detector.

The shape of a cocoa bean is outlined by putting a threshold to the peak-to-peak information. In this case, we used a threshold of 1.8 V peak-to-peak, meaning any pixel point with peak-to-peak higher than 1.8 V is considered as the plastic film reference, while the remaining pixel points are the cocoa bean points. This threshold was chosen by using the ptp intensity when the signal is transmitted through plastic film reference itself.

We further calculated the mean and standard deviation of the images for both peak-to-peak and time delay as specific quantitative classification parameters. The mean, μ , was calculated by averaging pixels' values of all the beans as:

$$\mu_{\text{time}} = \frac{\sum \Delta t_i}{N} \quad (3)$$

$$\mu_{\text{magnitude}} = \frac{\sum \text{ptp}_i}{N} \quad (4)$$

where Δt_i and ptp_i are the time-delay difference (in picoseconds) and peak-to-peak value (V) respectively of pixel i and N is the number of pixels. Similarly, the standard deviation (σ) of a sample was calculated as follow:

$$\sigma_{\text{time}} = \sqrt{\frac{(\Delta t_i - \mu_{\text{time}})^2}{N}} \quad (5)$$

$$\sigma_{\text{magnitude}} = \sqrt{\frac{(\text{ptp}_i - \mu_{\text{magnitude}})^2}{N}} \quad (6)$$

183

184 where μ is mean of the N sample population.

185 To compare the tolerance of the standard deviation to the mean of each sample, T parameter
186 can be calculated for both time delay and peak-to-peak as:

$$T_{\text{time}} = \sigma_{\text{time}} / \mu_{\text{time}} \times 100 \% \quad (7)$$

$$T_{\text{magnitude}} = \sigma_{\text{magnitude}} / \mu_{\text{magnitude}} \times 100 \% \quad (8)$$

189 The tolerance T is dependent on both σ and μ , therefore we can representatively use σ and T in data
190 analysis for qualification of the samples.

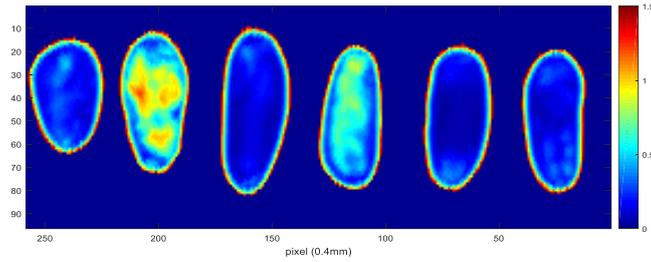
191

192 **3. Results and discussions**

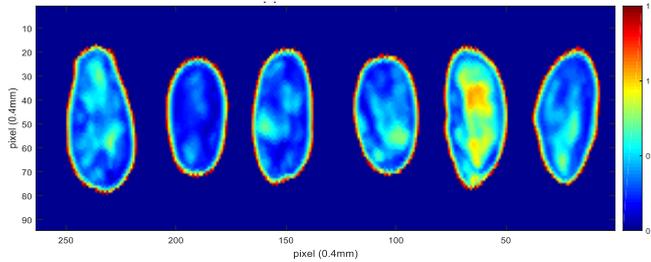
193 *3.1. Analysis of Samples in SET1*

194 The available twenty-two cocoa bean samples of SET1 were first measured and analysed for
195 the time-delay difference and peak-to-peak magnitude at each pixel. Pixels with peak-to-peak
196 value higher than the pre-determined threshold for our measuring system (1.8V) were from
197 the plastic film reference and are, then, referenced to zero. Figure 3 shows the images of the
198 peak-to-peak amplitude of the 22 cocoa beans. The 22 samples are placed from top to bottom
199 in the same order as in Table 1 (i.e., 11 pairs from 1 to 11 containing each 2 semi-beans). A
200 clear qualitative difference was observed between the non-fermented samples (with 0 day of
201 fermentation time) compared to the other samples.

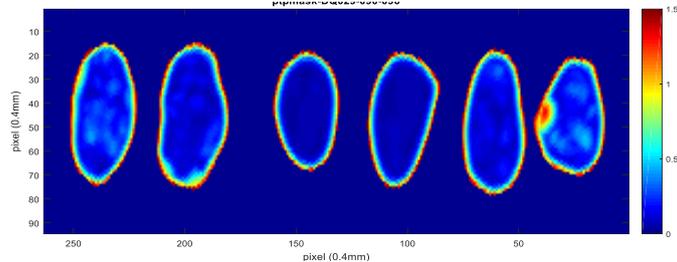
202



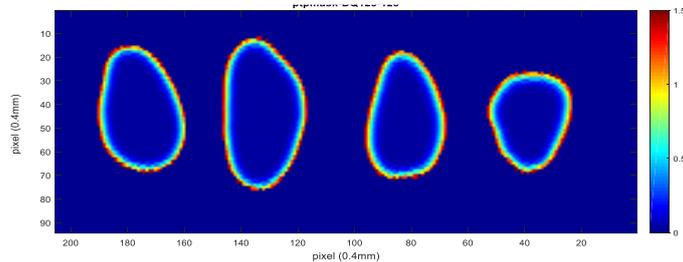
203



204



205

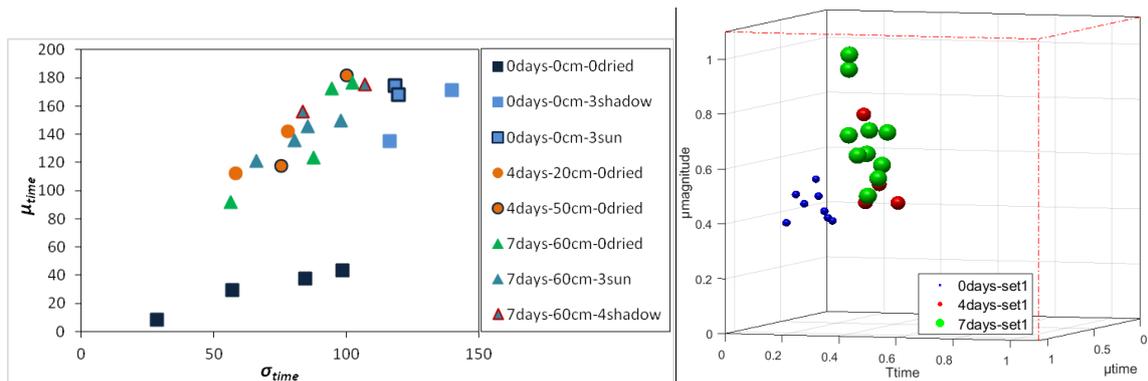


206 **Figure 3.** Images of time-delay difference (between the sample and the thin plastic film reference) of the 11
 207 pairs of samples.

208 The mean μ and standard deviation σ were then calculated for each cocoa bean sample. The
 209 results are plotted in Figure 4(a) showing μ_{time} vs. σ_{time} for group of samples collated together
 210 according to their fermentation time, fermentation height and drying time and method. A
 211 distinct difference was seen between the non-fermented samples (the squared dots) and those
 212 with different fermentation levels. The other samples of different fermentation heights (in
 213 cm) were not very well separated in this study. In particular, the non-fermented samples were

214 divided in two groups apparently according to the drying process, i.e., not dried versus dried
 215 samples (with sun or shadow).

216 It may be noted that the non-fermented and non-dried samples were clearly distinct from the
 217 remaining samples. On the other hand, the fermented samples (4 or 7 days) were clustered
 218 together without any differentiation. As time is reciprocal of frequency, the time parameters
 219 such as μ_{time} and σ_{time} are related to physical and chemical characteristics in THz frequency
 220 range of the samples. Therefore, this result may be linked to the strong absorption peak of
 221 NH_3 in the THz range. However, for further research and development, both experiment and
 222 data analysis will need to be done for better understanding the mechanisms influencing the
 223 THz results, especially if it is related to the fermentation heights of the samples.

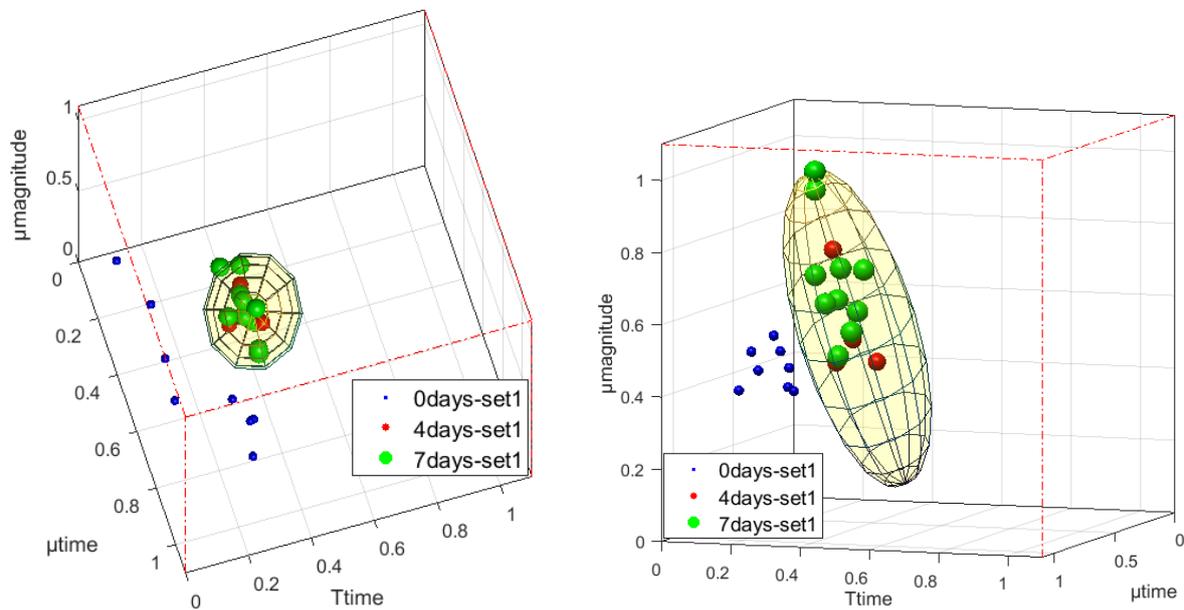


224 **Figure 4.** SET1: (a) Calculated μ_{time} vs. σ_{time} of the samples for different fermentation conditions (time in days,
 225 height in cm, and dried day and method); (b) 3D plot of normalised μ_{time} vs. T_{time} vs. $\mu_{magnitude}$.

227

228 For this study, we focus on qualifying the most important fermentation condition which is the
 229 fermentation time. Figure 4(b) plot shows a 3D view of the indices μ_{time} vs. T_{time} and vs.
 230 $\mu_{magnitude}$ of Set1 samples in which the samples were grouped based on their fermentation time
 231 (in days). To demonstrate fairly the influence of fermentation time to the measured THz
 232 indexes, the parameters were normalized before plotting. It showed a distinct separation in
 233 the second dimension (T_{time}) of non-fermented (0 day) vs. the fermented samples (4 or 7

234 days). In the following analysis, the samples were grouped based on the fermentation time
 235 only.



236

237 **Figure 5.** Comparison of samples obtained from SET1 using normalised parameters (μ_{time} vs. T_{time} vs. $\mu_{\text{magnitude}}$)
 238 and a fitting ellipsoid covering the fermented samples.0

239

240 In Figure 5, an ellipsoid fitting was calculated to cover all the fermented samples. If we use
 241 this ellipsoid as a determined threshold to separate fermented and non-fermented samples, the
 242 prediction was 100% matching to all the samples used in SET1.

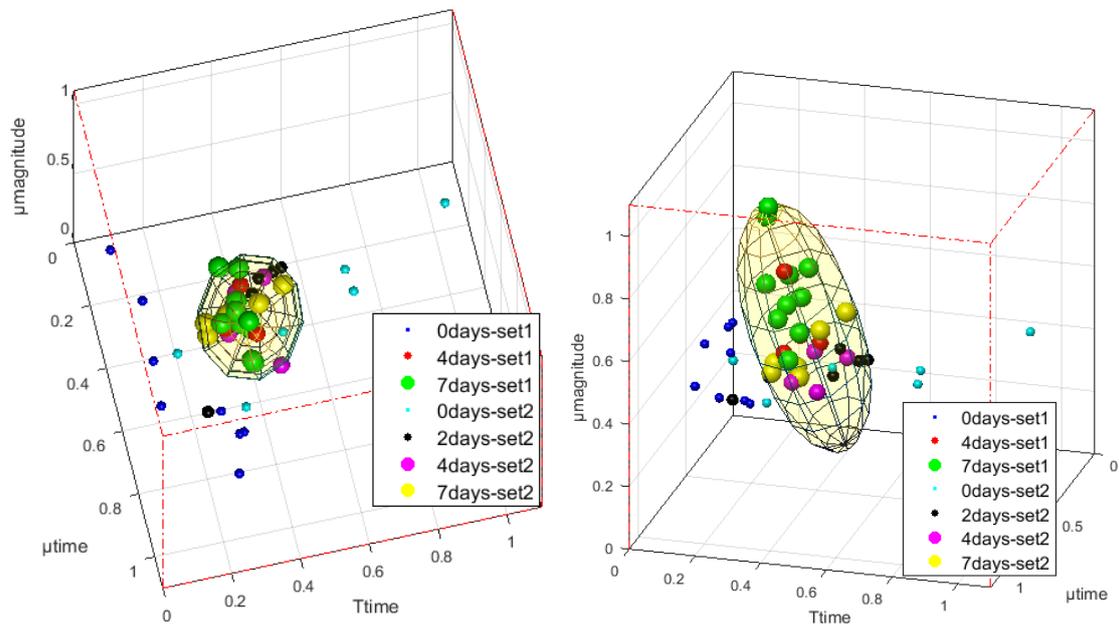
243

244 3.2. Analysis of second sets of samples (SET2) used for validation

245 To further validate the results obtained from SET1 samples and to understand the relation
 246 between different THz indexes and fermentation time, a second set (validation set) was
 247 measured and analysed. This validation set (SET2) consists of 22 different cocoa semi-bean
 248 samples as per Table 2.

249 Figure 6 compares SET1 and SET2 results. The same ellipsoid threshold calculated with
 250 SET1 was used to qualify the samples in SET2 for validation. Using the predefined standard,

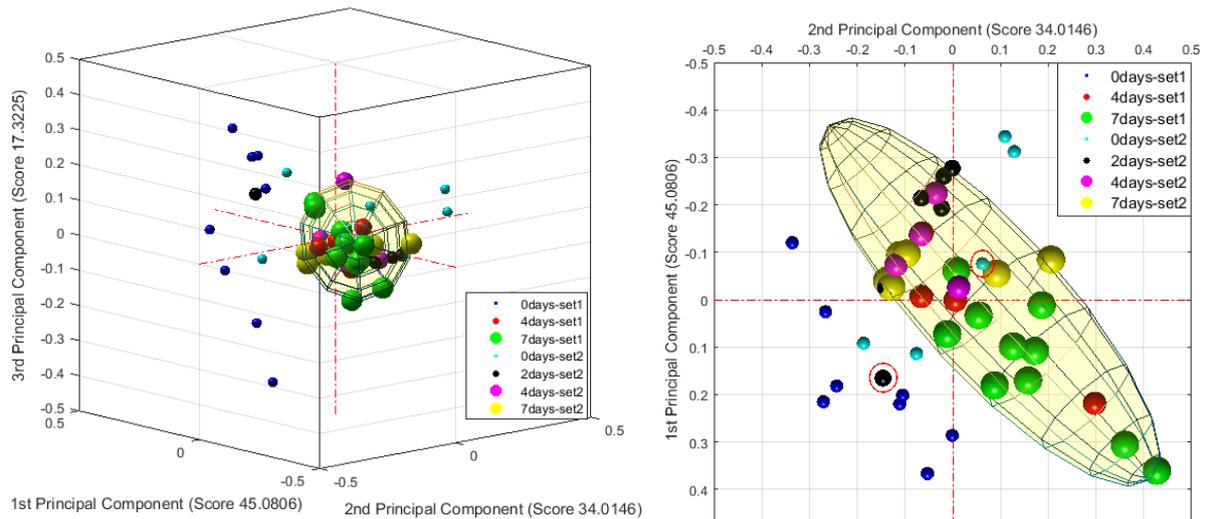
251 two mismatching cases were detected which are circled in red (one 2 days fermented sample
 252 is in non-fermented zone and one non-fermented sample is in the fermented zone).



253

254 **Figure 6.** Comparison of SET1 vs. SET2: 3D plot of normalised μ_{time} vs. T_{time} vs. $\mu_{magnitude}$ and fitting ellipsoid
 255 threshold of fermented samples in SET1.

256 To further investigate the data in 4 dimensions (μ_{time} , T_{time} , $\mu_{magnitude}$ and $T_{magnitude}$) principal
 257 components analysis (PCA) was deployed. The PCA was applied to all four parameters by
 258 normalising the parameters prior to applying the PCA. Figure 7 shows 3D plot revealing the
 259 scores of the first three principal components and a fitting ellipsoid for fermented samples.
 260 The first three components scored 96.3% which could be highly representing the data. From
 261 Figure 7 one can see that the fermented samples tend to be centred inside the ellipsoid
 262 whereas the non-fermented samples remain isolated and farther from the central region.
 263 There are two mismatching cases circled in red. The PCA calculation shows the same results
 264 as the study with three parameters (μ_{time} vs. T_{time} vs. $\mu_{magnitude}$).



265

266

267 **Figure 7.** Principal components analysis applied to Set1 and Set2: 3D plots of the first 3 principal components268 for analysis applied on 4 parameters: μ_{time} , T_{time} , $\mu_{magnitude}$ and $T_{magnitude}$ and a fitting ellipsoid covering the

269 fermented samples. The two samples marked in red circles are mismatching cases.

270

271 From Figure 7 one can see that samples of the same fermentation time grouped together (0, 2,

272 4 and 7 days) represented a unique cluster. In both measurements, non-fermented samples

273 remained identifiable distinctly compared to the fermented samples. Samples with a high

274 fermentation level occupied more centred distribution whereas non-fermented beans showed

275 a large, scattered divergence away from the centre.

276 Finally, using the same ellipsoid threshold, we can correctly predict 20 out of 22 samples for

277 SET2, giving 90.9% matching cases (the 2 samples in marked red circles are mismatching

278 cases). It is obvious that higher the population (number of samples), the better thresholds can

279 be calculated and the nearer we can reach to the true percentage of matching cases. For

280 example, if we count both SET1 and SET2 for the calculation, we correctly predicted 42 out

281 of 44 samples, making it 95.45% matching cases. As first feasibility study, it was

282 successfully demonstrated in this work that the use of THz technology for quality control of

283 cocoa beans can become a robust game changing technology offering better agility.

284 Improvement in confidence level is possible by further optimisation of the technique which
285 will be pursued in the follow-on work. This would particularly be beneficial and crucial to
286 apply to the emerging nanomanufacturing problems (Fan et al., 2021; Khatri et al., 2020;
287 Kumar Mishra et al., 2021)

288 **4. Conclusions**

289 A new measurement system was developed to characterise the fermentation levels in the
290 cocoa beans with an improved agility than the currently used methods used in this sector. Our
291 novel method involving the use of THz hyperspectral imaging showed new prospects for
292 deploying this technique as an industrial practice as a key non-destructive technique for
293 improved quality assurance for the chocolate business. Measurements involving two different
294 datasets obtained at different time periods helped to gain the confidence in the data which
295 proved to be a strong validation of the key results obtained and reported here. In this study, it
296 took about 4 minutes per sample of cocoa bean for its characterisation, but this time window
297 can further be squeezed by optimisation of the technique. The results show a distinct
298 difference between fermented and non-fermented samples, with prediction accuracy of 95%
299 while testing 44 samples.

300 The work has opened new possibilities in the arena of non-destructive and agile measurement
301 techniques to reliably characterise industrial crops, cereals and many other products used
302 commonly in the food sector.

303 **Acknowledgement**

304 This work was done under Tera4All project. We express great thanks to the Fédération
305 Wallonie Investissent and En Mieux for their support. Additional thanks to BELCOBRA
306 (Brazil-Belgium cooperation CAPES-WBI/2017-2019/361394) for financing the project:
307 ‘Optimization and implementation of analytical methods for the traceability and authenticity
308 of cocoa and chocolate’ from which the samples were prepared. This project is coordinated

309 by CRA-W and UFPA. Authors are grateful to CEPLAC (Medicilândia, PA, Brazil) for the
 310 mature fruits, to Teixeira L.E.O. and Freitas A.L. (UFPA) for their technical assistance for
 311 fermentation and drying process, and to Quentin Arnould (CRA-W) for laboratory support.
 312 SG greatly acknowledge the financial support provided by the UKRI via Grants No.
 313 EP/L016567/1, EP/S013652/1, EP/S036180/1, EP/T001100/1 and EP/T024607/1 and from
 314 the TFIN+ Feasibility study award to LSBU (EP/V026402/1), the Royal Academy of
 315 Engineering via Grants No. IAPP18-19\295 and TSP1332, EURAMET EMPIR A185 (2018),
 316 the EU Cost Action (CA15102, CA18125, CA18224 and CA16235) and the Newton
 317 Fellowship award from the Royal Society (NIF\R1\191571).

318

319 **References**

- 320 Assefzadeh, M.M., Jamali, B., Gluszek, A.K., Hudzikowski, A.J., Wojtas, J., Tittel, F.K., Babakhani, A., 2016.
 321 Terahertz trace gas spectroscopy based on a fully-electronic frequency-comb radiating array in silicon, 2016
 322 Conference on Lasers and Electro-Optics (CLEO). IEEE, pp. 1-2.
- 323 Barbin, D.F., Maciel, L.F., Bazoni, C.H.V., da Silva Ribeiro, M., Carvalho, R.D.S., da Silva Bispo, E., Miranda,
 324 M.d.P.S., Hirooka, E.Y.J.J.o.f.s., technology, 2018. Classification and compositional characterization of different
 325 varieties of cocoa beans by near infrared spectroscopy and multivariate statistical analyses. 55, 2457-2466.
- 326 Beans, C., 2015. Cocoa Beans: Chocolate & Cocoa Industry Quality Requirements.
- 327 Beg, M.S., Ahmad, S., Jan, K., Bashir, K.J.T.i.f.s., technology, 2017. Status, supply chain and processing of cocoa-
 328 A review. 66, 108-116.
- 329 Caporaso, N., Whitworth, M.B., Fowler, M.S., Fisk, I.D.J.F.c., 2018. Hyperspectral imaging for non-destructive
 330 prediction of fermentation index, polyphenol content and antioxidant activity in single cocoa beans. 258, 343-
 331 351.
- 332 Conway, E.J., Flood, J.C.J.B.J., 1936. An absorption apparatus for the micro-determination of certain volatile
 333 substances: The micro-determination of bromide, with application to blood and urine and observations on the
 334 normal human subject. 30, 716-727.
- 335 Cruz-Tirado, J., Pierna, J.A.F., Rogez, H., Barbin, D.F., Baeten, V.J.F.C., 2020. Authentication of cocoa
 336 (Theobroma cacao) bean hybrids by NIR-hyperspectral imaging and chemometrics. 118, 107445.
- 337 Fan, P., Goel, S., Luo, X., Upadhyaya, H.M.J.N., Metrology, 2021. Atomic-Scale Friction Studies on Single-Crystal
 338 Gallium Arsenide Using Atomic Force Microscope and Molecular Dynamics Simulation. 1-11.
- 339 Hindle, F., Kuuliala, L., Mouelhi, M., Cuisset, A., Bray, C., Vanwolleghem, M., Devlieghere, F., Mouret, G.,
 340 Bocquet, R.J.A., 2018. Monitoring of food spoilage by high resolution THz analysis. 143, 5536-5544.
- 341 Hue, C., Gunata, Z., Bergounhou, A., Assemet, S., Boulanger, R., Sauvage, F.-X., Davrieux, F.J.F.c., 2014. Near
 342 infrared spectroscopy as a new tool to determine cocoa fermentation levels through ammonia nitrogen
 343 quantification. 148, 240-245.
- 344 Ilgantileke, S., WAHYUDI, T., BAILON, M.G.J.J.o.f.q., 1991. Assessment methodology to predict quality of
 345 cocoa beans for export. 14, 481-496.
- 346 Khatri, N., Barkachary, B., Muneeswaran, B., Al-Sayegh, R., Goel, S., 2020. Surface Defects incorporated
 347 Diamond Machining of Silicon. International Journal of Extreme Manufacturing.
- 348 Kumar Mishra, R., Goel, S., Yazdani Nezhad, H., 2021. Computational prediction of electrical and thermal
 349 properties of graphene and BaTiO₃ reinforced epoxy nanocomposites. Biomaterials and Polymers Horizon 1, 1-
 350 14.

351 Nguyen, D., Dortu, F., Dispa, A., Hubert, P., Ziemons, E., Hernandez, Y., 2018. Terahertz hyper-spectral imaging
352 of lab-prepared versus commercial paracetamol tablets and potential applications, Unconventional Optical
353 Imaging. International Society for Optics and Photonics, p. 106773I.

354 Quelal-Vásquez, M.A., Lerma-García, M.J., Pérez-Esteve, É., Arnau-Bonachera, A., Barat, J.M., Talens, P.J.F.C.,
355 2019. Fast detection of cocoa shell in cocoa powders by near infrared spectroscopy and multivariate analysis.
356 99, 68-72.

357 Quelal-Vásquez, M.A., Pérez-Esteve, É., Arnau-Bonachera, A., Barat, J.M., Talens, P.J.F.C., 2018. Rapid fraud
358 detection of cocoa powder with carob flour using near infrared spectroscopy. 92, 183-189.

359 Santos, F.A., Palmeira, E.S., Jesus, G.J.D.i.b., 2019. An image dataset of cut-test-classified cocoa beans. 24,
360 103916.

361 Sunoj, S., Igathinathane, C., Visvanathan, R.J.C., Agriculture, E.i., 2016. Nondestructive determination of cocoa
362 bean quality using FT-NIR spectroscopy. 124, 234-242.

363 Teye, E., Huang, X., Sam-Amoah, L.K., Takrama, J., Boison, D., Botchway, F., Kumi, F.J.F.c., 2015. Estimating
364 cocoa bean parameters by FT-NIRS and chemometrics analysis. 176, 403-410.

365 Weiller, S., Tanabe, T., Oyama, Y.J.W.J.o.E., Technology, 2018. Terahertz non-contact monitoring of cocoa
366 butter in chocolate. 6, 268.

367