Vol.67: e24240166, 2024
https://doi.org/10.1590/1678-4324-2024240166
ISSN 1678-4324 Online Edition

BABT
Brazilian Archives of Biology and Technology

Article - Engineering, Technology and Techniques

Specific Dimensions of Soybean Grains Through Digital Images

Diana Soares Magalhães[1*]
https://orcid.org/0000-0001-5998-6837

Luís César da Silva[1]
https://orcid.org/0000-0001-9582-3998

Paulo Cesar Corrêa[1]
https://orcid.org/0000-0001-9843-1455

Camilla Sena da Silva[2]
https://orcid.org/0000-0003-0845-5010

Augusto Cesar Laviola de Oliveira[1]
https://orcid.org/0000-0002-1200-2551

[1]Universidade federal de Viçosa, Departamento de Engenharia Agrícola, Viçosa, Minas Gerais, Brasil; [2]Universidade Federal de Viçosa, Departamento de Agronomia, Viçosa, Minas Gerais, Brasil.

Editor-in-Chief: Alexandre Rasi Aoki
Associate Editor: Fabio Alessandro Guerra

Received: 20-Feb-2024; Accepted: 13-May-2024.

*Correspondence: dianasoaresmagalhaes@gmail.com; Tel.: +55-31-38745674 (D.S.M.)

HIGHLIGHTS

- Use of digital images to determine the dimensional properties of soybeans.
- Alternative method for obtaining grain dimensions.
- Engineering characteristics to obtain characteristic dimensions.
- The model proved to be efficient in determining the physical properties of soybeans.

Abstract: The determination of the physical properties of grains is laborious work and subject to subjectivities. Therefore, the objective of this work was to develop an automated method to measure the characteristic dimensions of soybean grains through digital images. Images of whole grains were captured and through them, the width, length, projected area, and perimeter of 50 grains were analyzed Using a computer application using Python programming language. The measurements for the grains obtained by the application were compared with those determined using a digital caliper. A Bland-Altman test was applied for the variables of width, length, and circularity, in which satisfactory results were obtained for the variable length, in which the method of obtaining measurements by digital images was not statistically different from the caliper method, making them equivalents. The width and circularity variables did not show equivalence, however, most of the data are within the 95% confidence interval. The variables of projected area and perimeter and the agreement between the data were analyzed employing an ANOVA followed by the Tukey test at 5% probability. The projected area variable did not differ significantly according to the methods used in the experiment, while the perimeter differed only for the method in which OpenCV library functions were used.

Keywords: automation; OpenCV; python; physical properties.

INTRODUCTION

Soybean (Glycine max), besides its great economic importance, is one of the most widespread crops in the world. Every year, records of soybean production and yield are broken to meet the growing demand of the increasing requirements of the market.

Knowledge of the physical properties of agricultural products is essential for carrying out activities such as characterization of the dimensions and shapes of the grains, description of the drying kinetics, the definition of parameters for quality monitoring, and sizing of dryers, conveyors, and equipment for cleaning and, or, grain classification [1, 2, 3]. Regarding the characterization of the dimensions and shapes of the grains, procedures are carried out aimed at measuring the dimensions of length, width, and thickness and calculating the circularity, perimeter, and projected area.

The procedures for determining the physical properties of grains and seeds are usually manual. Besides, they are laborious and time-consuming processes, subject to subjective interpretations by the operator in addition to impacting decision-making regarding processing routines and marketing practices. To speed up the process with reliability, Digital Image Processing (PDI) can be used, which consists of obtaining, processing, and analyzing digital images using properly verified and validated mathematical models [4, 5].

The use of computational tools and image processing has been increasing due to the reduction of equipment costs, the growing computational power of the hardware, and the interest in fast and non-destructive solutions [6]. Moreover, [4] state that the use of these technologies in the characterization of grains and seeds is due to the improvement of image processing techniques and the use of neural network modeling using machine learning. This innovative technique presents a novel avenue for researchers and practitioners in the field of agricultural science.

This study aims to contribute to the advancement of scientific knowledge through the application of PDI to determine the characteristic dimensions of soybean grains and explore the potential of digital images process as a practical and efficient tool for characterizing grains. Given the importance of knowing the dimensions of grains and the difficulty of the procedure, the objective of this work was to determine the characteristic dimensions of soybeans through digital images and image processing and to compare them with the dimensions obtained by a digital caliper.

MATERIAL AND METHODS

An algorithm capable of classifying and determining the physical properties of generic grains was developed. For the present work, soybeans were used as a case study to validate the applied methodology. Samples of fresh soybean from storage units located in the Center-West region of Brazil were used.

The digital images of the samples were obtained using a prototype made in MDF (Medium Density Fiberboard) in black with the dimensions length, height, and width equal to 0.20 m. A hole with a quadratic section of 2.5 x 2.5 cm was inserted on the upper face to fix an 8-megapixel resolution camera module V2 NoIR and next to it, the Raspberry Pi 4 Model B microprocessor with 4-GB of RAM was installed. Internally, to standardize the light intensity, white LED light strips were installed on the sides and top of the compartment.

A drawer is located in the lower part of the prototype. The samples of soybeans were stored in this drawer for image recording, thus forming the field of analysis, with length and width of 15.0 and 10.0 cm, respectively. The field of analysis was coated with satin vinyl foam (SVF) in blue, to avoid the movement of grains during the opening and closing of the drawer, to reduce the reflection caused by the internal lighting system, and to facilitate the segmentation of the images [7].

To activate the camera, a Python code was structured with the inclusion of the “picamera” and “time” modules. In this code, the following configuration parameter that was best adapted to the capture environment were defined, namely: rotation (180), resolution (3280 x 2464 pixels), framerate (30 frames per second), brightness (45), sharpness (100), saturation (10), contrast (20) and iso = 10. Those parameters were defined to obtain uniform images, to facilitate the segmentation process.

To determine the image spatial resolution, a coin with known dimensions was used. The coin had its images obtained in the same conditions used to capture the images of soybeans. Thus, by comparing the values of length, width, area, and perimeter obtained in the device with those found employing a caliper and the appropriate calculations, the conversion factors were determined.

To implement the algorithm, the Python programming language version 3.8 was used using the Jupyter Notebook™ development environment, as well as the OpenCV and Numpy libraries.

Images of whole grains were captured in the RGB pattern (Red, Green, and Blue) and, in sequence, transformed into the HSV (Hue, Saturation, Value) pattern to facilitate segmentation [8]. By doing so, the relationship between color and luminance is reduced, which leads to fewer influences from shaded areas or areas with greater light incidence. Consequently, there is greater precision in determining the dimensions and shape of the grains.

During the segmentation process, the H and V bands were added to facilitate the binarization, therefore, the Otsu Threshould method available in the OpenCV library was used. The binarization step consists of converting the images into two colors, black (digit 0) and white (digit 1). The image background (analysis field) was classified as 1 (white) and the grains as 0 (black). In this way, the dimensions and shape of the grains were obtained [9]). Before the binarization process, a medium blur filter (medianblur) was applied to soften noise in the images and minimize problems arising from shadows, dirt, and/or other flaws.

The findContours function available in the OpenCV module was applied in a loop with the objective of locating, delimiting, and counting the shape of the captured images of each grain. The functions cv2.contourArea, cv2.arcLength and cv2.boxPoints (rect) were used to determine the projected area, perimeter, width and length dimensions of each grain, respectively. Also, as a function of the area and radius found by the OpenCV functions, the circularity (C) was calculated using Equation 1.

$$C = \frac{4 * \text{area}}{\pi * (2 * radius)^2}$$

On what:
C = circularity, %;
area = shape area, mm²;
radius = shape radius, mm.

The computational algorithm also determines the longest axis of length and width of the shape of each grain. For comparative purposes, the area and perimeter were calculated using the length (a) and width (b) axes found by the computational algorithm using equations 2 and 3, respectively [10, 11, 12].

$$\text{Ap} = \frac{\text{ab}\pi}{4}$$

$$\text{P} = \frac{2\pi\ (\text{a} + \text{b}\ )}{4}$$

On what;
a = dimension of the longest taxis (length), mm;
b = dimension of the mean axis (width), mm;
Ap = projected area, mm²;
P = perimeter, mm.

To validate the procedures for determination of the dimensions and calculations of circularity, projected area, and perimeter, a comparison between these values and those obtained through laboratory routines was done [10, 11, 12].

The dimensions of length (a) and width (b) for 50 soybean grains were obtained using a Mitutoyo caliper, with a precision of 0.01 mm, by measuring the orthogonal axes of length (a) and width (b) as indicated in Figure 1. Each axis was measured in triplicate and the mean was calculated in order to reduce operator error. Through the measured axes, it was possible to calculate the circularity (C), projected area (Ap), and perimeter (P) through equations 02 and 03.

The same grains were submitted to the computational application, where it was possible to obtain their characteristic dimensions using digital image processing. Validation took place by comparing the results found through the processing of digital images and those obtained from measurements performed using the caliper.

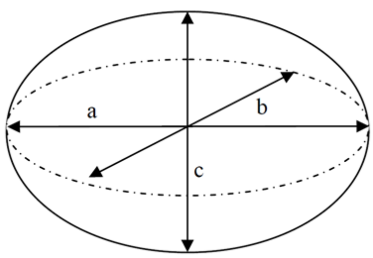


Figure 1. Triaxial spheroid with its respective characteristic dimensions: (a) length, (b) width and (c) thickness. Source: Corrêa and Goneli (2008).

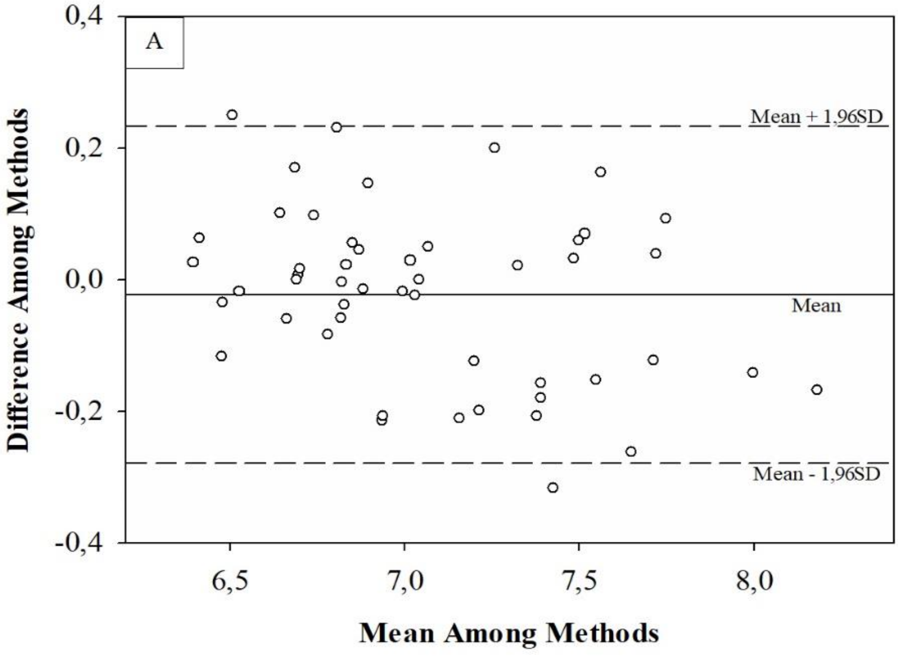
The use of IBM SPSS Statistic program [13] employed the Bland-Altman test [14] in order to analyze the relationship between the methods for obtaining characteristic dimensions obtained by (i) caliper and (ii) dimensions using the application for the variables of width (a), length (b) and circularity (C). The Bland-Altman test enables the checking of the dispersion between the differences between two variables and the means obtained between the methods used. Currently, it is the most suitable test for the analysis of agreement between methods [15, 16]. First, the T-test was performed to assess whether the differences between the methods are significantly different from zero, then the scatter plot was plotted in which it was possible to analyze data variability in relation to the mean of the differences and between the 95% confidence limits. Finally, a simple linear regression was carried out to assess the existence of proportion bias, which indicates the occurrence of a tendency for the data to vary below or above the mean.

The projected area and perimeter variables were calculated according to three methods in which the grain dimensions were obtained through (i) Caliper, (ii) Computer application, and (iii) OpenCV library functions (cv2.contourArea for area and cv2. arcLength for perimeter). In the evaluation of the methods, analysis of variance (ANOVA) and the comparison of means were performed according to Tukey's test at 5% of significance.

RESULTS AND DISCUSSION

The use of the T-test at a significance level of 5% ($p = 0.05$) enabled us to assess the methods applied to determine the width, length, and circularity, which corresponds to the use of (i) Caliper and (ii) computer application. According to the statistical analyses, among the methods used to determine the length (a) of soybean grains, a p-value of 0.230 was obtained, that is, greater than 0.05; therefore, there is no significant difference between the methods. Regarding the determination of the width (b) the p-value < 0.001, indicating that there is a statistical difference between the methods. As the circularity is the ratio between width and length (b/a), the same behavior as that obtained for b was found, where the T-test indicated the non-correlation between the methods for this variable (p-value < 0.001).

Figure 2 shows the Bland-Altman graphs, regarding the differences between methods for measuring length (a), width (b), and circularity (C). It is observed for length (a), the occurrence of greater accuracy for values greater than the mean, while for those below the mean, less dispersion is observed. Out of the 50 calculated values, an outlier was observed when considering the 95% confidence interval (Figure 2A). The Bland-Altman plot for the width variable (Figure 2B) shows greater data dispersion and the presence of three outliers. It is observed that for this variable, the computer application overestimated the data values, resulting in more negative differences between the methods. The mean difference was -0.10 mm ± 0.11 with a maximum difference of -0.31 mm, so it is important to emphasize that the differences found between the methodologies may refer to random or non-random errors caused in the acquisition of dimensions per caliper as well as during digital image processing [17]. Figure 2C shows the Bland-Altman graphic representation for the circularity variable (C), in which can be seen the same behavior observed for the width variable as well as the presence of outliers.



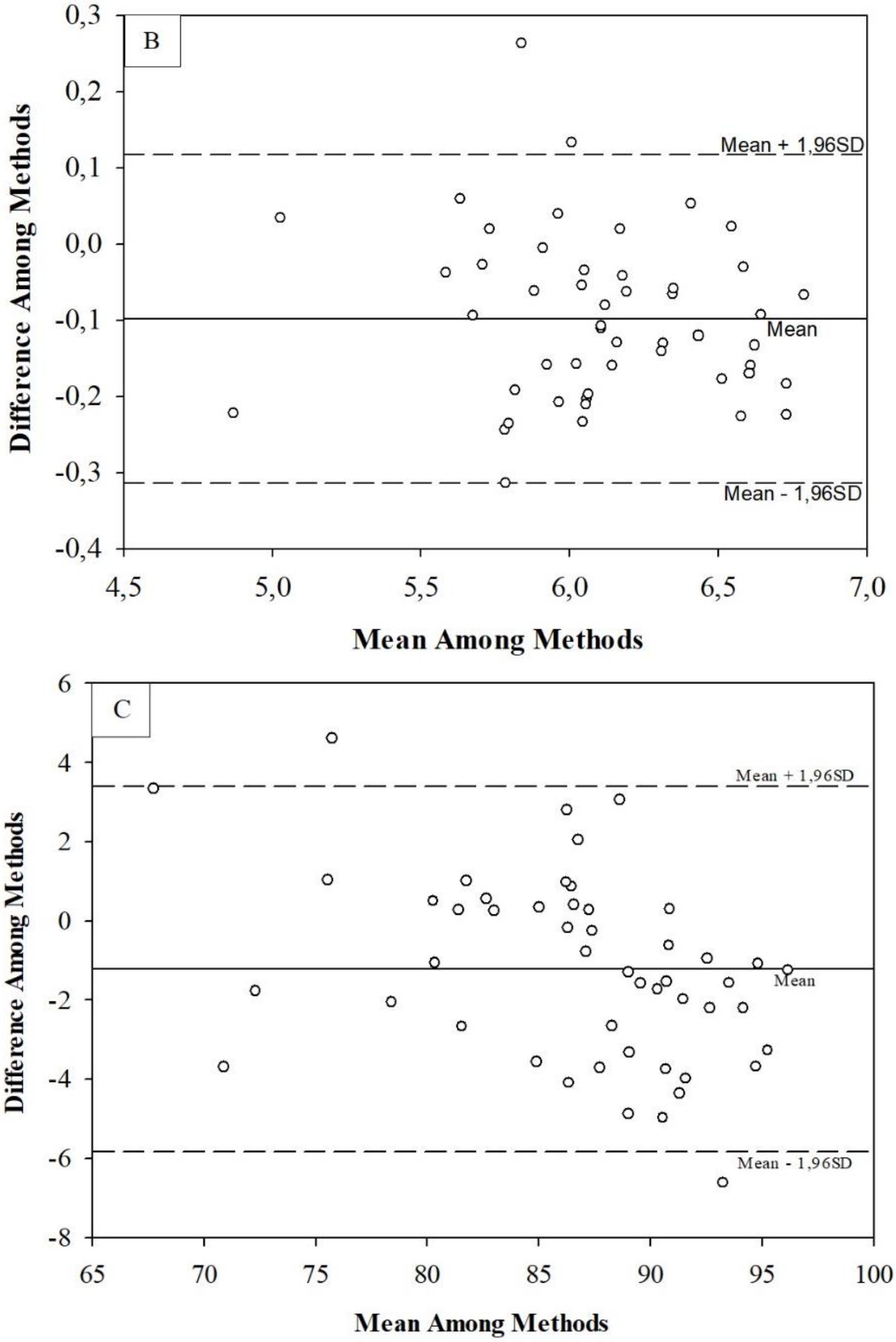


Figure 2. Bland-Altman plot for differences among measurement methods: length (A), width (B) and circularity (C).

The ANOVA results for the differences between the determination methods are shown in Table 1. The length determination methods do not show a significant difference between the data (F ≥0.05), so it can be said that there are no trends in underestimating or overestimating values, so the methods are concordant and equivalent. Concerning width, despite the methods being statistically different, ANOVA indicates that there is no bias, so there is no tendency to underestimate or overestimate values in relation to the mean. The analysis of variance for circularity pointed to a proportion bias demonstrating a tendency to overestimate the data.

Table 1. Summary of ANOVA for the differences among measurement methods.

SV	gl	Sum of the squares	Mean square	F
		Length (a)		
Regression	1	0.048	0.048	0.057
Residue	48	0.612	0.013	
Total	49	0.660		
		Width (b)		
Regression	1	0.008	0.008	0.425
Residue	48	0.586	0.012	
Total	49	0.594		
		Circularity (C)		
Regression	1	45.72	45.72	0.003
Residue	48	226.12	4.71	
Total	49	271.84		

Table 2 presents the summary of the Analysis of Variance (ANOVA) and the Coefficient of Variation (CV) for the projected area and perimeter variables according to parameters determined using: (i) measurements using caliper, (ii) characteristic dimensions per application, and (iii) OpenCV functions. It is observed that only for the perimeter variable, a significant F was obtained at the level of 1% of probability, indicating that at least one of the means is statistically different.

The means obtained for each method are shown in Table 3. It is observed that for both variables, the methods that use the OpenCV functions obtained higher means than the other methods. [18] when studying the use of OpenCV to determine leaf area in the field by digital images, observed that the image resolution directly affects the accuracy of the observed measurements. Therefore, one of the possible reasons for the differences found between the means may be related to the low resolution of the Pi Camera capture sensor used in this study.

Table 2. Summary of ANOVA for the projected area and perimeter of soybean grain.

		Mean square	
SV	**gl**	**Area**	**Perimeter**
Methods	2	22.61	32.79**
Residue	147	11.89	1.17
CV (%)	-	10.34	5.11

** F significant at 1% of probability.

Table 3. Means of the areas and perimeters for the respective methods.

Methods	Area	Perimeter
Caliper	32.51 a	20.60 a
Dimension per application	33.61 a	20.79 a
OpenCV function	33.73 a	22.09 b

*Means followed by the same letter in the column are not statically different at the probability level of 5% by the test of Tukey.

CONCLUSION

Considering the evaluations performed in this work, it can be concluded that it was possible to develop and evaluate a computer program capable of determining the physical properties of soybeans. In general, the model was efficient in determining the physical properties of soybeans. However, it is necessary to investigate ways to improve the capture method in order to reduce dimension errors for width (b). Image processing technologies is a fundamental process to automate physical properties determination processes.

Acknowledgments: This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.
Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

1. Goneli ALD, Corrêa PC, Magalhães FEA, Baptestini FM. [Volumetric Contraction and Shape of Castor Beans During Drying]. Acta Sci. Agron., 2011;33(1):1-8.
2. Ribeiro VS, Sobral MC, Almeida MM, Silva GF. [Physical Properties of Agricultural Products]. Rev. Bras. Prod. Agroind., Campina Grande, 2002; 4(1):1-6.
3. Brooker DB, Bakker-Arkema FW, Hall CW [Drying and storage of grains and oilseeds] Springer, New York, USA, 1992.
4. Jitanan S, Chimlek,P [Quality grading of soybean seeds using image analysis]. Int. J. Electr. Comput. Eng., 2019 9(5):3495-503.
5. Yang X, Hong H, You Z, Cheng F [Spectral and image integrated analysis of hyperspectral data for waxy corn seed variety classification] Sensors, 2015;15(7):15578- 94.
6. Mahajan S, DAS A, Sardana HK [Image acquisition techniques for assessment of legume quality] Trends in Food Science & Technology, 2015; 42(2):116-33.
7. Medeiros AD, Capobiango NP, Silva JM, Silva LJ [Interactive machine learning for soybean seed and seedling quality classification] Scientific reports, 2020;10(1):1-10.
8. Koschan A, Abidi M [Digital color image processing]. John Wiley & Sons, 2008.
9. Granitto PM, Navone HD, Verdes PF, Ceccatto HÁ [Weed seeds identification by machine vision]. Comput. Electron. Agric., 2002;33:91-103.
10. Wu Ang, Zhu J, Yang Y, Liu X, Wang X, Wang L, et al. [Classification of corn kernels grades using image analysis and support vector machine]. Adv. Mech. Eng., 2018;10(12):1687814018817642.
11. Mohsenin NN [Physical properties of plant and animal materials]. Gordon and Breach Sci. Pub, 1972, p. 734.
12. Sakai N, Yonekawa S [Three-dimensional image analysis of the shape of soybean seed]. J. Food Eng, 1992;15(3):221-34.
13. IBM, Statistics. SPSS Statistics. Available from: https://www.ibm.com/products/spss-statistics.
14. Bland JM, Altman DG [Comparing methods of measurement: why plotting difference against standard method is misleading]. The Lancet, 1995;346(8982):1085-87.
15. Doğan NÖ [Bland-Altman analysis: A paradigm to understand correlation and agreement]. Turk. J. Emerg. Med., 2018;18(4):139-41.
16. Hirakata VN, Camey S [Concordance Analysis Between Bland-Altman Methods]. Clin. Biomed. Res, 2009;29(3).
17. Araujo MEV, Barbosa E, Oliveira ACL, Milagres RS, Pinto FAC, Correa PC [Physical properties of yellow passion fruit seeds (*Passiflora edulis*) during the drying process]. Scientia Horticulturae, 2020;261:109032.
18. Hariadi TK, Fadholi F, Chamim ANN, Utama NA, Prabasari I, Riyadi S [Development of Leaf Area Meter Using Open CV for Smartphone Application]. Telkomnika, 2018;16(4):1857-63.