



OPEN The use of image analysis to study the effect of moisture content on the physical properties of grains

Łukasz Gierz^{1✉}, Mustafa Ahmed Jalal Al-Sammarraie², Osman Özbek³ & Piotr Markowski⁴

Designing machines and equipment for post-harvest operations of agricultural products requires information about their physical properties. The aim of the work was to evaluate the possibility of introducing a new approach to predict the moisture content in bean and corn seeds based on measuring their dimensions using image analysis using artificial neural networks (ANN). Experimental tests were carried out at three levels of wet basis moisture content of seeds: 9, 13 and 17%. The analysis of the results showed a direct relationship between the wet basis moisture content and the main dimensions of the seeds. Based on the statistical analysis of the seed material, it was shown that the characteristics examined have a normal or close to normal distribution, and the seed material used in the investigation is representative. Furthermore, the use of artificial neural networks to predict the wet basis moisture content of seeds based on changes in their dimensions has an efficiency of 82%. The results obtained from the method used in this work are very promising for predicting the moisture content.

Keywords Physical properties, Image analysis, Moisture content, Overall dimensions, Artificial neural network

Corn (*Zea*) is classified as one of the three most important cereal crops in the world. For humanity, of the 6 species¹ of corn, the most important in economic terms is common corn (*Zea mays L.*), also called the type species². The grains (seeds) of individual corn varieties differ in size, color, shape, and texture. The most common varieties are yellow, red, and white corn kernels, which constitute an important traditional food, animal feed, and industrial raw material that play an important role in agricultural production^{3,4}. In turn, beans belong to the legume family. In the world, it is produced and consumed mainly in the form of dry beans and green beans. Legumes are the main source of proteins and nutrients such as oil, fiber, starch, vitamins, and minerals^{5,6}. Corn is grown in more than 170 regions of the world, with the main growing regions being North America, Asia, and South America. Corn production in the United States (USA), China, Brazil and Argentina accounts for up to 64.63% of world production⁷. Meanwhile, world bean production reached Asia, Africa and the Americas (North, Central, and South) with 49.72%, 24.40%, and 24.36% respectively⁸. Sustainable food systems are required to ensure sustainable food production that can meet the growing demand for a world population expected to reach 9.8 billion people by 2050⁹. In today's competitive market, manufacturers must offer products classified by physical characteristics such as appearance, size, color, and internal health status. Moreover, identifying the varieties of grains helps farmers in selecting the right grains to sow and marketing because the grains they sell thus sell will be of the basic standard of marketers¹⁰. The physical properties of grains are important for equipment design and analysis of product behavior during agricultural processes such as planting, harvesting, threshing, cleaning, classification, and processing. The main grain dimensions are useful for selecting screen spacers and calculating the grinding force (reduction) when reducing size. They can also be used to calculate the surface area and size of grains, which is important when modeling grain drying, ventilation, heating and cooling processes¹¹. The manual process of grain sorting is time-consuming and inefficient, especially in the case of large production volumes. For these reasons, new methods have been introduced, such as vision systems for classification, sorting, and measurement. Due to increasing consumer demand and technological progress, industries are engaged

¹Institute of Machine Design, Faculty of Mechanical Engineering, Poznan University of Technology, ul. Piotrowo 3, 60-965 Poznan, Poland. ²Department of Agricultural Machinery and Equipment, College of Agricultural Engineering Sciences, University of Baghdad, Baghdad, Iraq. ³Department of Agricultural Machineries and Technologies Engineering, Faculty of Agriculture, Selçuk University, 42250 Konya, Turkey. ⁴Department of Heavy-Duty Machines and Research Methodology, University of Warmia and Mazury, ul. Oczapowskiego 11, 10-719 Olsztyn, Poland. ✉email: lukasz.gierz@put.poznan.pl

in the automation of processing steps such as grading, sorting, packaging, determining the quality of fruits and vegetables, automation and design of handling mechanisms, processing, peeling, etc. Physical properties such as weight, volume, surface area, bulk density, etc. must be known, therefore, it is necessary to check the relationship between them¹². Alamery and Al-Badri¹³ developed mathematical models to predict grain mass and size based on grain length, width, and thickness. Image-processing techniques are used after agricultural produce is harvested to ensure food quality, hygiene, and processing suitable for human consumption.

The selected physical characteristics of seeds are an important parameter in the assessment of food quality. Consumers expect agricultural products of the same size and characteristic shape. Eboibi and Uguru¹⁴ used statistical analysis such as analysis of variance (ANOVA) and Duncan's multiple range test to investigate the physical properties of bean varieties. Yenge et al.¹⁵ determined the physical and mechanical properties of corn kernels as a function of moisture content using standard techniques. Kruszelnicka¹⁶ examined the physical and mechanical properties of corn grains from the point of view of designing mechanical processing systems. Today, the latest research on cereal products is carried out using machine vision systems and image processing technologies. Using the methods and equipment mentioned, products are examined for many physical properties, such as color¹⁷, texture, quality¹⁸, and size. Image processing uses various mathematical and numerical models to associate one or more properties and also to predict various physical properties of products. Several studies have been conducted on the mathematical and numerical relationship between the physical properties of fruits. Ghabel et al.¹⁹ describe the relationship between mass and geometric mean diameter. The surface area and volume of models with different shapes can be measured by estimating three mutually orthogonal axes²⁰. Grain classification has been studied by several researchers using image analysis. Shahin and Symons²¹ used color and grain size using a planar imaging system to determine color classification. In the case of Sicilian and Canadian lentil varieties, color characteristics were found to be good predictors for variety recognition. Kılıç et al.²² developed a bean classification system based on seed size and color. The results showed that this method allowed for the correct classification of seeds. Guevara-Hernandez and Gil²³ developed an automated vision system for classifying wheat and barley grains. The results of verification of the developed system showed that accuracy exceeding 99% can be achieved when extracting morphological, color, and textural features from grains. Tahir et al.²⁴ successfully used machine learning algorithms to extract features from images of western Canadian durum wheat, western Canadian red spring wheat, and barley to investigate the effect of moisture content on the ability to classify the physical properties of the imaged grains. Wilson et al.²⁵ measured the grain size distribution of wheat starch using image analysis and laser diffraction techniques. Arasan et al.²⁶ used image analysis methods to determine the grain size distribution (granulometric composition). The literature also indicates that the physical properties of cereal grains depend mainly on the water content²⁷. Research on predicting the impact of moisture content based on the physical properties of grains (seeds) using image processing technology has not been precisely defined; therefore, there is a gap in the state of knowledge in this area, which was decided to fill. The article sets out three specific goals. The first was to determine the effect of humidity on the physical properties of beans and corn (geometric dimensions), such as length, width, and thickness. The second was to assess the usefulness of the developed method using the image analysis technique to determine the geometric dimensions of the grain. The third goal was to use an artificial neural network (ANN) to predict the moisture content of beans and corn grains based on their geometric dimensions.

Materials and methods

Sample preparation

It should be noted that for the purposes of the investigation, the authors used varieties of crop plants sown in Poland. It is also worth noting that, based on statistical data from 2021, Poland became the third corn producer in the European Union, after France and Romania, with a 11% share in EU production (<https://www.fao.org/faostat/en>). The research material consisted of beans of the Eureka variety (*Phaseolus coccineos L.*) and corn seeds of the Gramatura variety (*Zea Mays L.*) All research material came from PlantiCo Zielonki Sp. z o. o. Spółna, Breeding and Seed Plant in Nochowo, located in Greater Poland Province. Based on the announcement of the Marshal of the Sejm of the Republic of Poland about the Legal Protection of Plant Varieties of January 22, 2021 (Journal of Laws of 2021, item 213) and the breeder's declaration that the indicated varieties: Eureka (beans), Grammaturgia (corn) are protected by law by the breeder; the authors have received this permission. The breeder agreed to provide the above-mentioned plant material, and it's use is in compliance with the national guidelines of the Main Seed Warehouse, a production plant in Nochowo. With the consent of the breeder, the authors could use the obtained plant material only for scientific research purposes, including research on, among others, oversize dimensions.

The samples were manually cleaned to remove foreign bodies, dust, dirt, broken and immature grains. The weight of one sample was 500 g, which was determined using an electronic scale RADWAG PS 4500/X with an accuracy of 0.01 g. The initial moisture of the samples, which was 9%, was determined using a digital grain moisture meter AR991 with a measurement range of 7.5–50% and measurement accuracy of 0.1%, made in China. The grains were subjected to another measurement procedure after moistening to the level of 13% and 17% by adding measured amounts of filtered water M_w according to (Eq. 1), also used in other works^{28,29}.

$$M_w = \frac{(W_2 - W_1)}{100 - W_2} \cdot M_p \quad (1)$$

where: M_w —the amount of water needed to increase the moisture content (g), M_p —the weight of the sample (g), W_1 —the initial moisture content of the sample (%), W_2 —the assumed moisture content of the sample (%).

After adding a certain amount of water, the seeds were closed in polyester string bags and stored at (5 °C) in the refrigerator for 7 days to obtain an even distribution of moisture in the seeds. The humidity stabilization time was determined on the basis of studies conducted previously^{30–33}, in which it was 5 or 15 days. Then the

required amounts of seeds were taken from string bags to perform measurements, and the remaining seeds were used to control the wet basis moisture content.

The amount of moisture in a biological material can be expressed in terms of wet weight (wet weight) or dry weight (dry weight), expressed as a decimal or percentage³⁴. For the purposes of this research, the wet basis moisture content expressed as a percentage of M_{wb} was determined based on Eq. (2)³⁴, defined as the mass of moisture contained in the seeds per unit mass of undried seeds:

$$M_{wb} = \frac{W_o - W_d}{W_o} \cdot 100 \quad (2)$$

where: M_{wb} —wet basis moisture content expressed as a percentage (%), W_o —initial weight of undried seeds (g), W_d —mass of dry seed (g).

System methodology

The research methodology consists of six steps: image capture of each grain, image processing, grain edge detection (find contours from image of each grain), data acquisition and description, data preservation, and prediction of moisture content on a wet basis.

The three main dimensions, length (L), width (W) and thickness (T), were measured for 100 randomly selected kidney beans and maize (A group of grains was taken from one sample and then it was divided again until we reached 100 grains) using the proposed system.

Preparation of materials for camera tests

Using the pixy2 camera, the main grain dimensions, such as length (L), width (W), and thickness (T), were measured for 100 randomly selected beans (seeds) of beans and corn. The pixy2 camera is characterized by low purchase cost and ease of use. The camera used contains an Omni Vision OV9715 sensor, 1280 × 800, 1/4. It can process 50 images per second (20 ms per image). It can be connected to a computer via the USB port. Additionally, it is characterized by a very low weight of 27 g and a low current consumption of 140 mA³⁵. The camera is connected to the Arduino Uno controller using the serial communication protocol, thanks to which the camera takes photos from different angles. The camera was successfully used in the analysis of fruit images³⁶. An LED camera light source was used, the camera light intensity of which is 341 Lux. The light intensity of the camera was measured using a digital lux meter. Measuring range 0.1 ~ 200,000 lx, 0.01 ~ 20,000 Fc, resolution: 0.1 lx/0.01 Fc, accuracy ± 4%. The camera was mounted vertically above the measuring table (grain). This mounting provided a field of view of 60 degrees horizontally, 40 degrees vertically from a height of 15 cm. Figure 1 shows the measurement system.

To obtain accurate measurements, an attempt was made to reduce the shadow from the program settings. The pixy2 camera was calibrated by entering images of different grains. After preparing the program with different grain images, the test environment was adjusted by monitoring the lighting intensity. The lighting intensity should be bright enough to allow photos to be taken with the required sharpness. The image capture unit was isolated from room light and sunlight, which had a positive effect on the calibration process³⁷. In each measurement cycle, the grains were arranged in three columns and six rows. During the measurement, a space was deliberately left between the grains because the grain arrangement greatly influences the overlap of the image frames, which affects the measurement accuracy. It has been proven that placing samples too close to each other, in our case grains (seeds), causes cross-contamination during measurement³⁸.

Grain dimensions measurement system

A computer vision system consists of two units: (1) an image processing unit and (2) a pattern recognition unit. The system is controlled by a microcontroller. The system captures an image of the grain (seed) and sends it to the image analysis processor, and the physical properties of the various grains are calculated. Figure 2 shows the system algorithm. The system code was written in C++.

To process the images recorded with the pixy2 camera, the PixyMon program was used. The program is easy to use and sends the necessary data to the control unit. To measure the main dimensions, the edge of the grain must be detected by detecting the edge pixels in the image. Edge pixels are a group of points (pixels) on a curve that separate adjacent points (pixels) or points (pixels) on the other side of the curve that differ in brightness³⁹. The beans were placed in front of the Pixy2 camera and the 'Set Caption' option was selected. This option allows you to select the desired grain displayed on the screen in a rectangle of the desired size. PixyMon then highlights the entire grain. PixyMon divides each seed into a "block" that assigns a signature. This block signature can be sent to the control unit (Arduino UNO), e.g. the height, width, and distance of the block. You need a computer to distinguish the blocks. These data are stored in the memory card of the Pixy2 camera after it is disconnected from the computer.

After the analysis, we receive the following data: Block (this is the position of the block in the table), sig (signature number 1-7), x and y (coordinates of the block in the visible frame), width, height (length), index (unique number assigned to the object on the display) and age (the number of frames in which a given block was viewed, 1-255).

The measurement dimensions of the pixy2 camera are similar to the human eye in that distant objects are perceived as having small dimensions and increase as the distance between the camera and the object shortens. Therefore, when accurately calculating the main dimensions, the distance of the camera from the grains is calibrated and the results obtained from the camera are compared with those obtained by manual calculation using a caliper. A caliper with an accuracy of 0.02 mm was used for manual measurement. After adjusting the

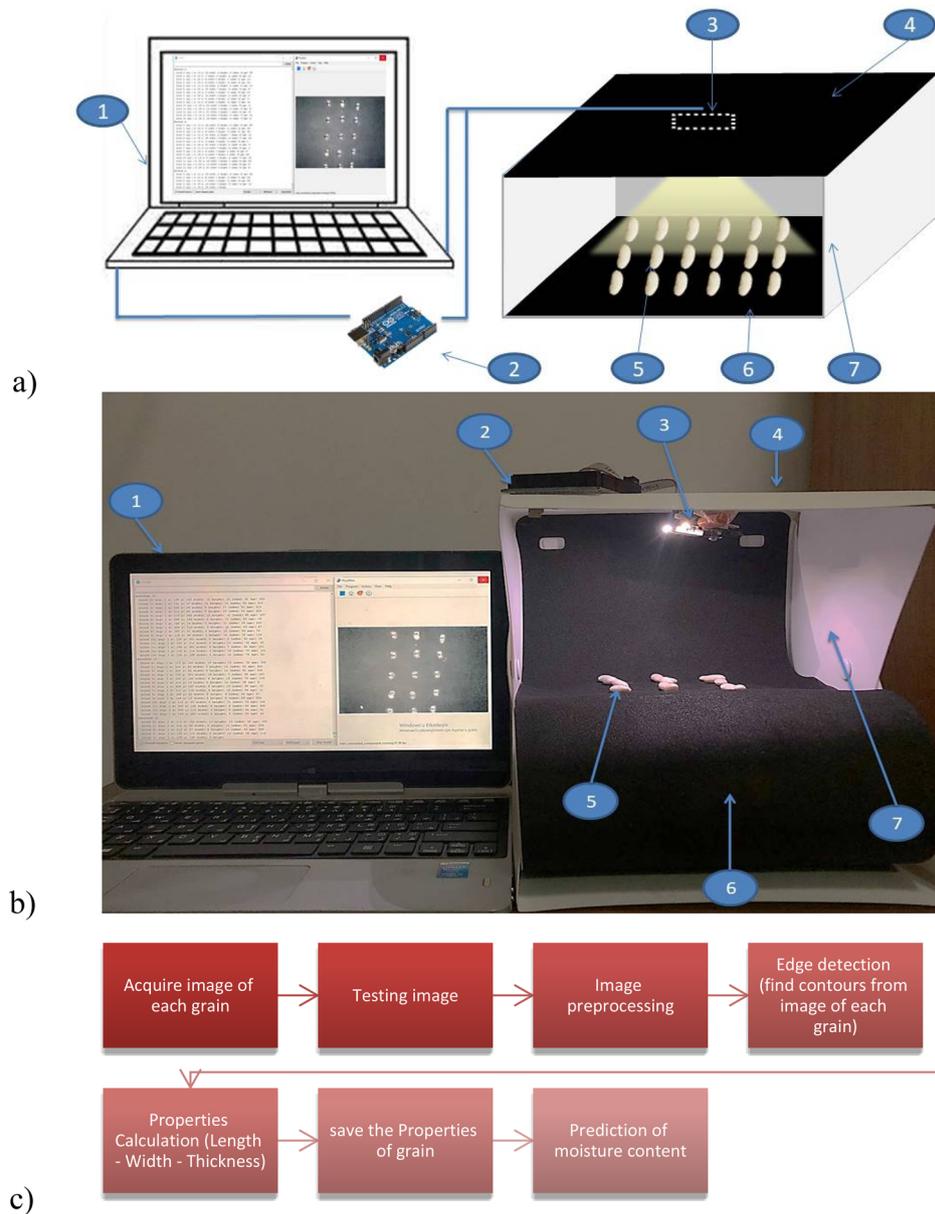


Figure 1. Measurement system, (a) schematic diagram: (1) laptop, (2) control unit (Arduino UNO), (3) pixy2 camera, (4) cover, (5) grains, (6) black plate, (7) grain chamber: (b) real view; (c) block diagram.

distance of the camera lens from the measuring table, the identified dimensions are divided by 10 to obtain greater measurement accuracy.

The structure of an artificial neural network

When creating artificial neural networks of the multilayer perceptron (MLP) type, the first step is to design training sets. To obtain appropriate results, it is very important to determine the characteristics being studied and the number of training instances. From the proposed system, grain size data were obtained, indicating the wet basis moisture content. An artificial neural network was designed, which includes an input, hidden, and output layer. The number of hidden layers, including the number of neurons and the type of activation function in the hidden and output layers, was determined using the Weka tool. The Weka tool includes a set of visualization tools and algorithms for data analysis and predictive modeling. This tool is publicly available and has easy access to user interfaces. The program is entirely implemented in the Java programming language, so it works on almost any modern computer platform. The neural network contained an input layer that contained the main dimensions of the grains (length, width, and thickness), two hidden layers for each variable with a different number of neurons (6.5), while the output layer contained the wet basis moisture content (see Fig. 3 for a diagram). The quality of the obtained networks was assessed based on the error of the training, testing, and verification sets, and then the root mean square error was determined. During the training process of each neural network, the data was

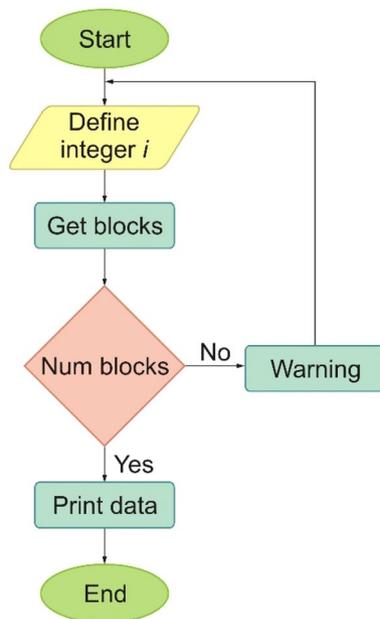


Figure 2. System algorithm.

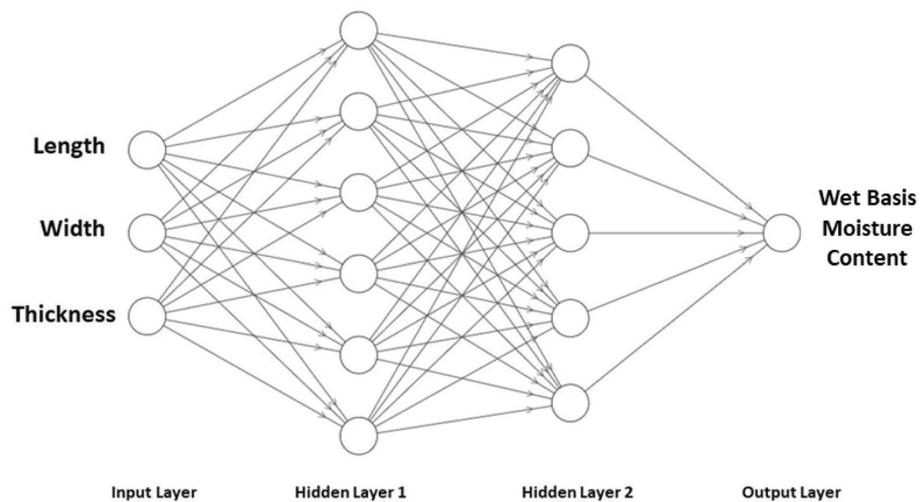


Figure 3. Neural network test.

randomly divided into three groups: training sets, test sets and validation sets. The training data set used in the learning process consisted of 80%, test sets 10%, and validation sets 10%. To define the performance measure, the definition defined in the previous work⁴⁰ was used.

TP rate: Which represent positive samples. The TP rate equation is:

$$TP / (TP1 + FN1) \tag{3}$$

FP rate: A positive error, which is considered a positive sample, but is actually a negative sample. The FP rate equation is:

$$FN1 / (TP1 + FN1) \tag{4}$$

Precision: The rate of correct cases for a given category, and correctly predicted by the classifier. The precision equation is:

$$TP1 / (TP1 + FP1) \tag{5}$$

Recall: The ratio of the truly integer number in the retrieval results in the integer number in the entire data set (recovered and unrecovered). The recall equation is:

$$TP1 / (TP1 + FN1) \quad (6)$$

F-measure: is the discrepancy between the precision P and R of the weighted mean harmonic index. The F-Measure equation is:

$$\frac{TP1 \cdot TN1 - FP1 \cdot FN1}{\sqrt{(TP1 + FP1)(TP1 + FN1)(TN1 + FP1)(TN1 + FN1)}} \quad (7)$$

ROC area: The area under the ROC curve (AUC) provides another way to evaluate the average performance of a model.

PRC Area: Use exact registration formulas to complement ROC equations to obtain the full spectrum during analysis and selection.

Accuracy: Correctly classified cases. The accuracy equation is:

$$\frac{TP1 + TN1}{TP1 + TN1 + FP1 + FN1} \quad (8)$$

where: TP1– positive true, FP1– positive false, TN1– negative true, FN1– negative false.

Methods of developing research findings

The research results were statistically processed using the following procedures:

1. Determining basic measures of position and dispersion of measurement results of individual physical properties of grain (length, width, and thickness) obtained using image analysis techniques and calipers;
2. Verification of hypotheses assuming that the empirical distributions of the values of individual geometric characteristics of beans and corn grains (measured with individual measuring instruments and dependent on humidity) are consistent with the normal distribution. Calculations were carried out using the following tests: W Shapiro–Wilk and Komogorow-Smirnow (with Lilliefors' correction) and χ^2 -Pearson;
3. Assessment of the homogeneity of variance of the measured feature (Levene's test) if its distribution is consistent with the normal distribution;
4. Comparison of the the significance of differences between the average values of the individual physical properties of beans and corn grain:
 - a) Measured with different t-Student measuring devices, when the measured features had a normal distribution and with the Mann–Whitney U test, when the distribution of features was not compliant with the normal distribution,—the assumption of the Student's t-test was not satisfied;
 - b) Depending on humidity; If the distribution of the measured characteristic was consistent with the normal distribution and there was homogeneity of variance, one-way analysis of variance (ANOVA) was used for the calculations. Otherwise, the nonparametric Kruskal–Wallis test was used. Furthermore, when statistically significant differences were demonstrated between the mean values of the measured trait, 'post-hoc' tests were performed. In the first situation, when it was concerned with ANOVA analysis, the significance of differences was checked with tests of different sensitivity, that is, Fisher, Scheffe, HSD Tukey and Duncan, and if it was related to the results of the Kruskal–Wallis test, the option of multiple comparisons was used for mean "rank" values. The purpose of these analyzes was to distinguish so-called homogeneous groups.

The calculations were performed at the significance level of $\alpha = 0.05$, using the STATISTICA PL statistical program.

Results and discussion

Dimensions of the main grains

Two methods were used in the investigation to measure the main dimensions of corn and bean grains: the first using the Pixy2 camera and the second using a caliper (control method). First, the distribution of individual grain dimensions (length, width, and thickness) measured with various devices was checked to be consistent with the normal distribution. Table 1 presents measures of the significance distribution of individual geometric features (length, width, and thickness) of the grain, which were measured with various measurement tools.

Taking into account the value (p) obtained at the hypothesis level of $p < \alpha = 0.05$, the H_0 should be rejected in favor of the alternative hypothesis H_1 , assuming that the distribution of a given feature is not consistent with the normal distribution. It was similar in the case of the the width of beans in the new method and for three levels of humidity and the thickness of the beans at 17% humidity (Table 1). Hypothesis H_1 , assuming the lack of compliance of the examined characteristic with the normal distribution, was also observed for the width of the corn grain at 13% and 17% humidity. In control tests, the distribution is abnormal for the width of the bean grain at 9% and 13% moisture and thickness at 13% and 17% moisture.

In the case of the remaining results on the geometric dimensions of beans and corn, there was no grounds to reject the null hypothesis H_0 . This is evidenced by the probability values (p) of the tests used (Shapiro–Wilk test and the K-S test with Lilliefors correction), which are statistically insignificant ($p > 0.05$).

If the distribution of the physical properties of beans and corn grains was consistent with the normal distribution, parametric tests were used to determine the significance of differences between the average values of these

Geometric dimension	Measurements made with the new method					Control measurements				
	N	maks D	Lillief. <i>p</i>	S-W W	<i>p</i>	N	maks D	Lillief. <i>p</i>	S-W W	<i>p</i>
Bean grain—moisture 9%										
Length	20	0.17.215	<i>p</i> < 0.15	0.93.983	0.23.797	20	0.15.571	<i>p</i> > 0.20	0.93.168	0.16.636
Width	20	0.19.195	<i>p</i> < 0.05	0.89.182	0.02.904	20	0.19.111	<i>p</i> < 0.05	0.89.360	0.03.134
Thickness	20	0.13.003	<i>p</i> > 0.20	0.95.132	0.38.754	20	0.13.720	<i>p</i> > 0.20	0.94.666	0.31.923
Bean grain—moisture 13%										
Length	20	0.15.786	<i>p</i> > 0.20	0.91.764	0.08.925	20	018.059	<i>p</i> < 0.10	0.94.929	0.35.638
Width	20	0.24.761	<i>p</i> < 0.01	0.81.674	0.00.155	20	0.28.897	<i>p</i> < 0.01	0.71.324	0.00.006
Thickness	20	0.16.805	<i>p</i> < 0.15	0.93.003	0.15.464	20	0.16.130	<i>p</i> < 0.20	0.90.130	0.04.361
Bean grain—moisture 17%										
Length	20	0.16.947	<i>p</i> < 0.15	0.92.596	0.12.910	20	0.18.269	<i>p</i> < 0.10	0.94.184	0.25.970
Width	20	0.29.726	<i>p</i> < 0.01	0.89.583	0.03.447	20	0.16.764	<i>p</i> < 0.15	0.92.676	0.13.372
Thickness	20	0.20.909	<i>p</i> < 0.05	0.88.719	0.02.388	20	0.20.312	<i>p</i> < 0.05	0.89.624	0.03.508
Maize grain—moisture 9%										
Length	20	0.11.296	<i>p</i> > 0.20	0.95.375	0.42.769	20	0.18.502	<i>p</i> < 0.10	0.93.522	0.19.446
Width	20	0.12.551	<i>p</i> > 0.20	0.96.673	0.68.487	20	0.25.266	<i>p</i> < 0.01	0.92.313	0.11.386
Thickness	20	0.22.525	<i>p</i> < 0.05	0.90.765	0.05.749	20	0.22.988	<i>p</i> < 0.05	0.92.565	0.12.732
Maize grain—moisture 13%										
Length	20	0.14.150	<i>p</i> < 0.10	0.95.595	0.13.045	20	0.15.089	<i>p</i> < 0.05	0.95.729	0.14.469
Width	20	0.18.366	<i>p</i> < 0.01	0.94.163	0.04.308	20	0.22.554	<i>p</i> < 0.01	0.92.432	0.01.186
Thickness	20	0.13.649	<i>p</i> < 0.10	0.95.986	0.17.644	20	0.13.219	<i>p</i> < 0.10	0.95.973	0.17.472
Maize grain—moisture 17%										
Length	20	0.17.668	<i>p</i> < 0.10	0.94.619	0.31.299	20	0.21.358	<i>p</i> < 0.05	0.91.851	0.09.274
Width	20	0.25.879	<i>p</i> < 0.01	0.85.961	0.00.776	20	0.30.226	<i>p</i> < 0.01	0.85.127	0.00.560
Thickness	20	0.17.205	<i>p</i> < 0.15	0.90.523	0.05.173	20	0.16.023	<i>p</i> < 0.20	0.94.396	0.28.457

Table 1. Parameters of significance distribution of individual geometric features (length, width, and thickness) of grain measured using various measuring instruments. Significant values are in [italics].

characteristics obtained using different measuring devices. However, if the basic condition for using parametric tests was not met, the nonparametric Kruskal–Wallis test was used.

The results of detailed comparative analyzes of the measurement results of the same physical properties of beans and corn, using different measuring instruments, for three levels of grain moisture, are presented in Figs. 4 and 5. As can be seen from the (*p*) value, there was no grounds to reject the hypothesis in any case H_0 assuming equal values of individual grain dimensions determined using different measuring instruments.

Figure 4 shows the results of the measurements of the physical properties of beans depending on their wet moisture for three different levels (9, 13, 17%), a new measurement method (1), and a caliper (2). The longest kernels (seeds) had dimensions of 9.38, 9.67, 6.25 mm and 9.33, 9.58, 6.20 mm, while the widths of the kernels (seeds) were 6.20, 6.25, 6.30 mm and 6.10, 6.20, 6.20 mm. The high moisture content in the wet state also influenced the grain thickness, as the two previous methods obtained grain thicknesses of 5.21, 5.23, 5.35 mm and 5.16, 5.17, 5.30 mm, respectively.

Figure 5 shows the results of the measurements of the physical properties of corn grains depending on their wet moisture for three different levels (9, 13, 17%), a new measurement method (1), and a caliper (2). The longest kernels (seeds) had dimensions of 7.38, 7.72, 8.12 mm and 7.32, 7.7, 8.09 mm, while the widths of the kernels (seeds) were 5.16, 3.80, 5.30 mm and 5.12, 3.70, 5.30 mm. The high moisture content in the wet state also influenced the grain thickness, as the two previous methods obtained grain thicknesses of 3.75, 3.80, 4.50 mm and 3.71, 3.76, 4.04 mm, respectively.

The new method for analyzing the basic grain dimensions developed in this work turned out to be as accurate as the traditional method based on caliper measurement; therefore, we relied solely on the results of the new method in further analysis.

The results of the analyses (ANOVA and Kruskal–Wallis) showed the appearance of statistically significant differences between the mean values of the individual dimensions of bean and corn grains depending on their wet basis moisture content ($p < \alpha = 0,05$)—these are the results marked in red in the table. This was confirmed by ‘post hoc’ research, which allowed us to isolate the so-called homogeneous groups, i.e. those whose mean values do not differ statistically significantly. It should be emphasized that although ‘post hoc’ tests with different “sensitivities” were used to assess the significance of differences between mean values (ANOVA), the final results of the calculations were the same (Table 2 shows sample calculations and the results of Duncan’s test).

The dimensional analysis of the grains showed that the higher their moisture content in the wet state, the larger their basic dimensions (length, width, and thickness). From the above results, it was concluded that there is a direct relationship between the moisture content and the main dimensions of the grains. Researchers achieved similar results in their articles^{41,42}.

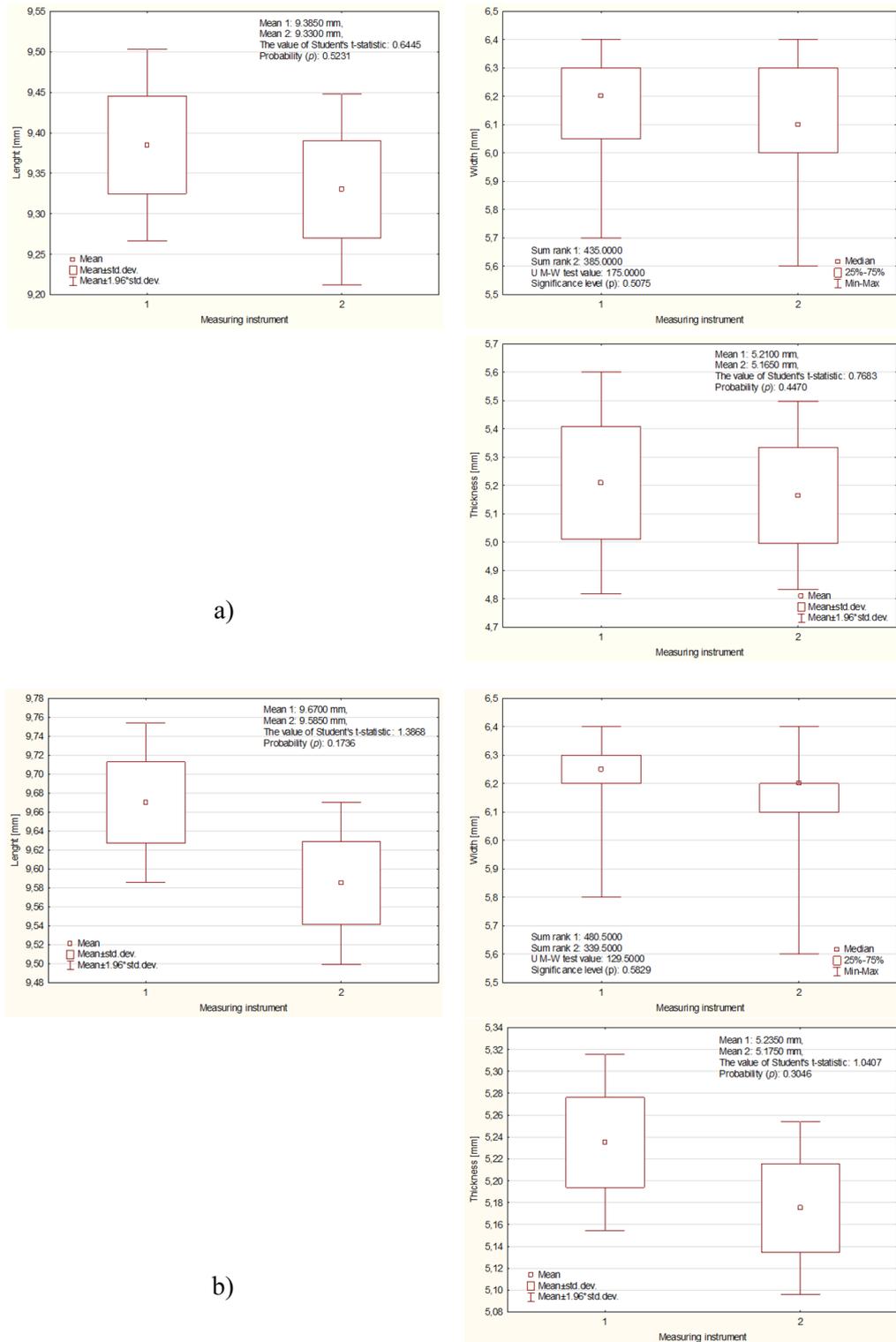


Figure 4. Mean and standard deviation values of fava beans at different moisture levels: (a) 9%, (b) 13%, (c) 17%.

Differences in the nature of changes in the dimensions of bean and corn grains may result from differences between individual grains in terms of water absorption capacity, as well as differences in their chemical composition and morphological structure. The ability of grains to absorb water is important not only during storage, but also during sorting and conditioning before grinding^{43–45}. Moisturizing the grains is a technological process that is frequently used. The wheat grain milling technology requires grain with a moisture content of 15.5–16.0%,

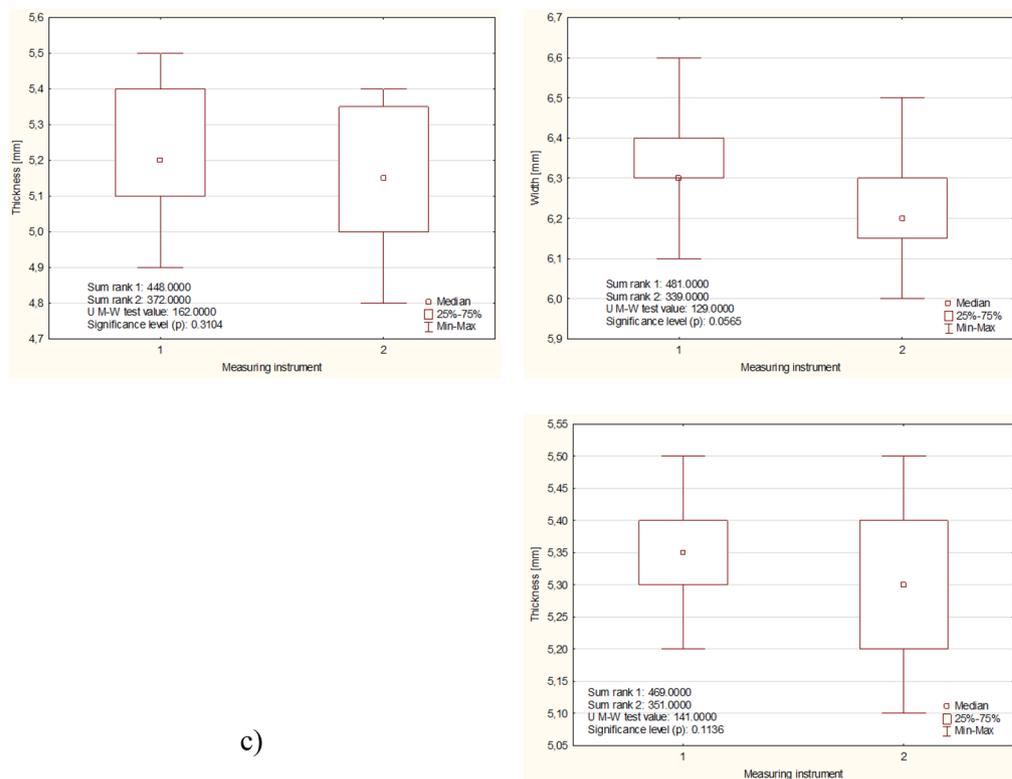


Figure 4. (continued)

with storage humidity ranging from 12 to 14%; therefore, in the milling industry, grains are often moistened to the indicated level⁴⁶.

Based on the measurement data, the average values of grain dimensions were compared with the two previous methods to verify the accuracy of the measurement. It was found that the new method of analyzing the main dimensions of grains is as accurate as the traditional method (measurement with a caliper); therefore, the new method can be used to measure the main dimensions of grains and thus predict the moisture content of selected grain species.

Artificial neural network performance

The ANN of the type of multilayer perceptron (MLP) was used to predict the wet basis moisture content of grains (seeds) based on the main dimensions (length, width, and thickness). The performance of the ANN configuration was evaluated several times using different datasets and configurations. It is best to configure an ANN using two hidden layers, with six cells in the first layer and five cells in the second layer. The correctly classified cases were 19 with a measurement accuracy of 79.2%, a mean absolute error of 0.18, and a mean squared error of 0.34. To determine the network performance measure, the evaluation measures used in the analysis are defined in detail and presented in Table 3.

It is clear from Table 3 that the mean positive samples (*TP Rate*) is 0.79. Where the percentage forecast at 9% humidity level gave the highest degree of 0.83. The average positive error (*FP rate*) is 0.11, with the smallest error occurring at 17% humidity. The average percentage of true-form prediction accuracy (*Precision*) was 0.82. Although the *F-Measure* was 0.79, the ROC area was 0.88 and the PRC area was 0.87.

To better understand the results, a confusion matrix was created that shows the prediction error (Fig. 6). Pink cells show incorrect predictions and blue cells show correct predictions. Where a grain moisture content of 9% is expected, a moisture content of 13% is also expected at one time. A grain moisture content of 13% was predicted to be twice that of 9%. It was also predicted that the grain with a moisture content of 17% would once have a moisture content of 9% and also once have a moisture content of 13%.

The reason for the overlapping results is that the recorded values are very close to each other, which classifies them into one group and, therefore, makes it difficult for the algorithm to predict the indicated humidity value. To illustrate this, the scatter plot shown in Fig. 7. Square-shaped points show results that were incorrectly predicted.

Based on the above results, the neural network to predict the moisture content. Previously, a similar network successfully predicted the sweet taste of oranges⁴⁷. The proposed network can also predict various human diseases^{48,49}. Going further, the studied network was used to predict soil penetration resistance^{50–52}, as well as soil pH⁵³. Neural networks were also used to predict the water quality index⁵⁴.

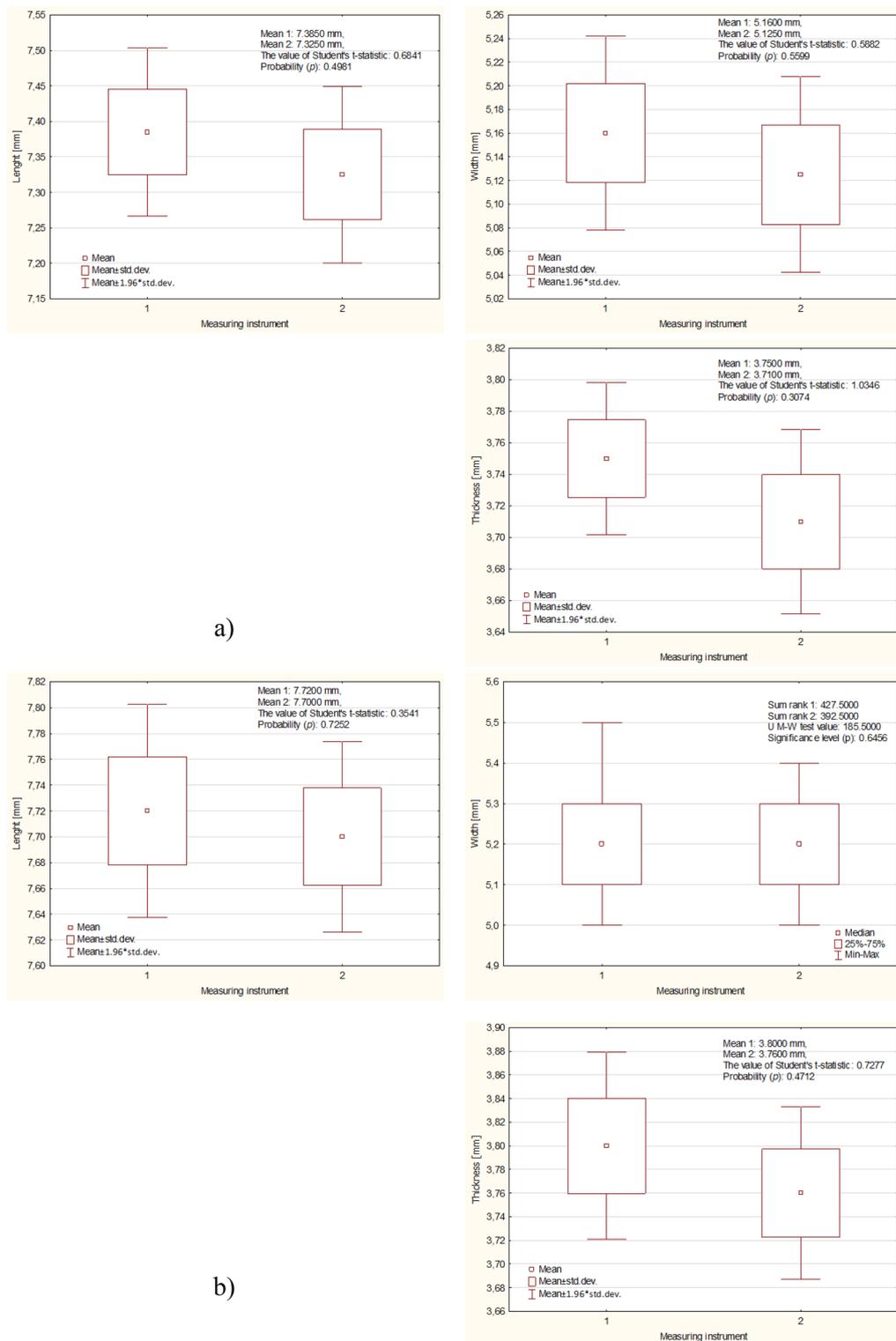


Figure 5. Mean and standard deviation values of fava maize at different moisture levels: (a) 9%, (b) 13%, (c) 17%.

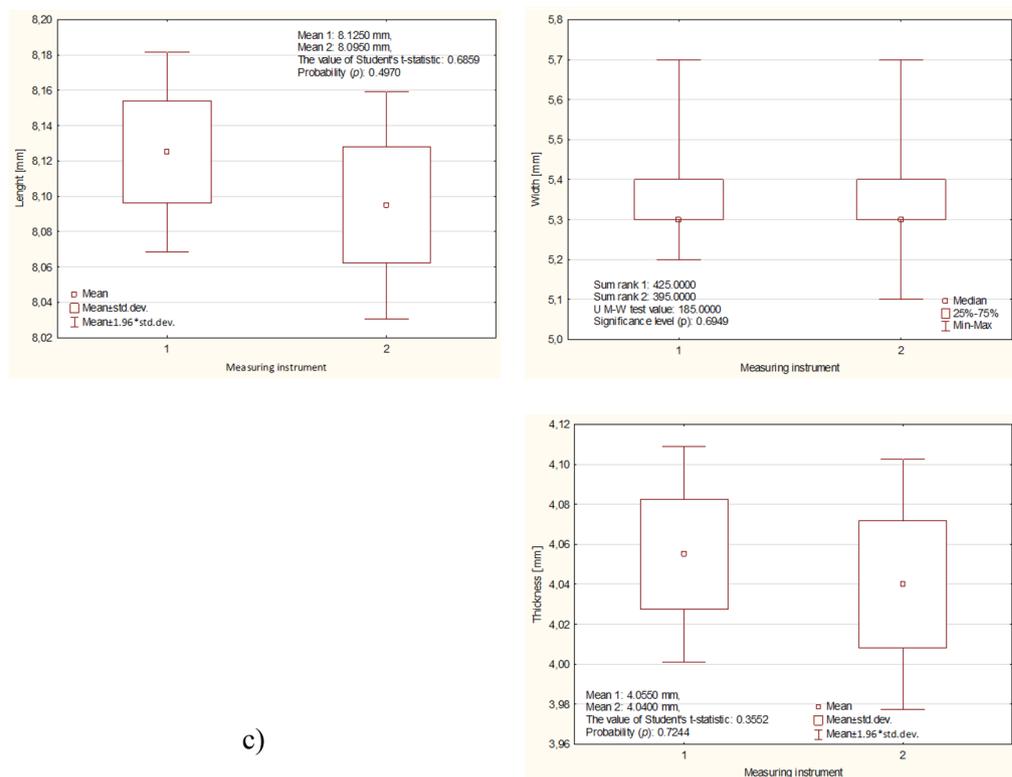


Figure 5. (continued)

Conclusions

In this work, a new system based on image processing technology was developed to measure the main grain dimensions (length, width, and thickness) to determine the effect of wet basis moisture content. The results obtained with the new method were compared with those obtained with the traditional method using a caliper. When examining the normal distribution of the results in the Shapiro–Wilk test for ($p = 0.05$), the hypothesis of normal distribution was accepted. The innovative approach to grain analysis developed in this work turned out to be as accurate as the traditional method based on measurements made with calipers. Based on the results obtained with the new method, it was found that there is a direct relationship between the wet basis moisture content and the main dimensions of the grains. Additionally, an artificial neural network of the multilayer perceptron (MLP) type was developed to predict the wet basis moisture content of grains based on the main dimensions of the grains. The algorithm consisted of three layers, an input layer representing the main dimensions of the grain, two hidden layers, and an output layer representing the wet basis moisture content of the grains. The first hidden layer consisted of 6 cells, while the second hidden layer consisted of 5 cells. Based on the test results, 19 cases were presented correctly classified with a measurement accuracy of 79.2%, the mean absolute error was 0.18, and the mean square error was 0.34. Furthermore, the proposed artificial neural network gave an efficiency of 82%. We conclude that, on the results of modern technological methods, it is possible to predict wet basis moisture content based on dimensional measurements using an artificial neural network. Due to this, it is possible to predict the dimensions of grains (seeds) based on wet basis moisture content measurement and correctly set the sowing dose, taking into account the volumetric nature of seed dosing using dispensers. Moreover, in further research, we plan, in addition to the use of artificial neural networks, to use other artificial intelligence techniques or hybrid models to predict the dimensions of grains (seeds) based on moisture measurements.

Bean grain length (Kruskala-Wallis test) H (2, N = 60) = 35.98052; p = 0.0000			
Variable –grain moisture [%]	Probability of multiple comparisons		
	Bean grain–moisture 9% R = 14.53	Bean grain–moisture 13% R = 29.55	Bean grain–moisture 17% R = 47.43
9	–	0.0196	0.0000
13	0.0196	–	0.0036
17	0.0000	0.0036	–
Bean grain width (ANOVA test) Efekt: F(1.056) = 2.2749; p = 0.1058			
Variable –grain moisture [%]	Probability Duncan's tests		
	Bean grain–moisture 9% X = 5.7933	Bean grain–moisture 13% X = 5.8700	Bean grain–moisture 17% X = 5.9800
9	–	0.3835	0.0437
13	0.3835	–	0.2112
17	0.0437	0.2112	–
Bean grain thickness (Kruskala-Wallis test) H (2, N = 60) = 6.832006; p = 0.0328			
Variable –seed moisture [%]	Probability of multiple comparisons		
	Bean grain–moisture 9% R = 25.60	Bean grain–moisture 13% R = 27.25	Bean grain–moisture 17% R = 38.65
9	–	1.0000	0.0544
13	1.0000	–	0.1170
17	0.0544	0.1170	–
Maize grain length (Kruskala-Wallis test) H (2, N = 60) = 42.15251; p = 0.0000			
Variable –grain moisture [%]	Probability of multiple comparisons		
	Maize grain–moisture 9% R = 13.75	Maize grain–moisture 13% R = 28.48	Maize grain–moisture 17% R = 49.28
9	–	0.0230	0.0000
13	0.0230	–	0.0005
17	0.0000	0.0005	–
Maize grain width (Kruskala-Wallis test) H (2, N = 60) = 16.23271; p = 0.0003			
Variable –grain moisture [%]	Probability Duncan's tests		
	Maize grain–moisture 9% X = 22.57500	Maize grain–moisture 13% X = 25.97500	Maize grain–moisture 17% X = 42.95000
9	–	1.0000	0.0007
13	1.0000	–	0.0063
17	0.0007	0.0063	–
Maize grain thickness (Kruskala-Wallis test) H (2, N = 60) = 29.40179; p = 0.0000			
Variable –grain moisture [%]	Probability of multiple comparisons		
	Maize grain–moisture 9% R = 19.550	Maize grain–moisture 13% R = 24.68	Maize grain–moisture 17% R = 47.28
9	–	1.0000	0.0000
13	1.0000	–	0.0001
17	0.0000	0.0001	–

Table 2. Summary of the calculations results that verify the significance of differences between mean values of bean and maize grins depending on their moisture content on a wet basis. X—mean value (mm); R—mean rank.

Moisture content on a wet basis %	TP rate	FP rate	Precision	Recall	F-measure	ROC area	PRC area
9	0.83	0.16	0.62	0.83	0.71	0.84	0.78
13	0.82	0.15	0.81	0.81	0.81	0.89	0.90
17	0.71	0.00	1.00	0.71	0.83	0.91	0.89
Weighted avg	0.79	0.11	0.82	0.79	0.79	0.88	0.87

Table 3. Results of network performance determination.

Classified as →		a	b	c
Low	a	5	1	0
Normal	b	2	9	0
High	c	1	1	5

Figure 6. Confusion matrix showing error in prediction.

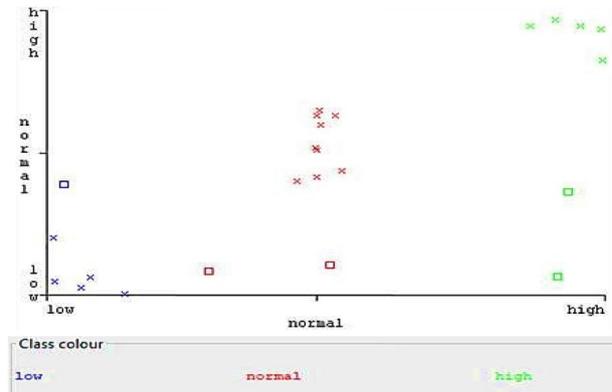


Figure 7. Scatter plot.

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Received: 18 November 2023; Accepted: 29 April 2024

Published online: 22 May 2024

References

1. *Plants of the World online*. Royal Botanic Gardens, Kew. <https://powo.science.kew.org/taxon/urn:lsid:ipni.org:names:19280-1> (access 2024-03-07).
2. *Index Nominum Genericorum* <https://naturalhistory2.si.edu/botany/ing/> (access 2024-03-08)
3. Wu, A. *et al.* Classification of corn kernels grades using image analysis and support vector machine. *Adv. Mech. Eng.* **10**, 168781401881764 (2018).
4. Al-Aridhee, A. H. A. & Mahdi, N. T. Influence of irrigation systems and cover crop on water productivity, and maize growth. *Iraqi J. Agric. Sci.* **53**, 1465–1475 (2022).
5. Laurent, B., Ousman, B., Dzudie, T., Carl, M. F. M. & Emmanuel, T. Digital camera images processing of hard-to-cook beans. *J. Eng. Technol. Res.* **2**, 177–188 (2010).
6. Al-Dulaimi, N. H. & Al-Jumaili, M. A. Role of iron, zinc and organic nutrient on yield of green beans. *Iraqi J. Agric. Sci.* **48**, 462 (2017).
7. Nchanji, E. B. & Lutomia, C. K. Regional impact of covid-19 on the production and food security of common bean smallholder farmers in Sub-Saharan Africa: Implication for sdg's. *Global Food Secur.* **29**, 100524 (2021).
8. Siddiq, M., Uebersax, M. A. & Siddiq, F. *Global Production, Trade, Processing and Nutritional Profile of Dry Beans and Other Pulses, Dry Beans and Pulses: Production, Processing, and Nutrition* 1–28 (Wiley, 2022).
9. Mergos, G. Population and food system sustainability. In *International Handbook of Population Policies* 131–155 (Springer, 2022).
10. Chen, X., Xun, Y., Li, W. & Zhang, J. Combining discriminant analysis and neural networks for corn variety identification. *Comput. Electron. Agric.* **71**, S48–S53 (2010).
11. Mohsenin, N. N. *Physical Properties of Plant and Animal Materials* (Gordon and Breach Science Publishers, 1970).
12. Shahir, S., Visvanathan, R., Nambi, V. E. & Chandrasekar, V. Mathematical modeling of physical properties of banana fruit for machine vision systems using image processing method. *JARDCS* **10**, 611–617 (2018).
13. Alamery, M.S. & Al-Badri, S.B.S., Determination the wheat weight and volume: Mathematical approach. *Proc. IOP Conference Series: Earth and Environmental Science*, **1213**, 012103 (2023).
14. Eboibi, O. & Uguru, H. Statistical analysis of the physical properties of varieties of beans (*Phaseolus vulgaris* L.) influenced by maturity stage. *Niger. J. Technol.* **37**(4), 1176–1184 (2018).
15. Yenge, G. B., Kad, V. P. & Nalawade, S. M. Physical properties of maize (*Zea mays* L.) grain. *J. Krishi Vigyan* **7**, 125–128 (2018).
16. Kruszelnicka, W. Study of selected physical-mechanical properties of corn grains important from the point of view of mechanical processing systems designing. *Materials* **14**(6), 1467 (2021).
17. Przybył, K. *et al.* Fruit powder analysis using machine learning based on color and ftr-atr spectroscopy—case study: Blackcurrant powders. *Appl. Sci.* **13**, 9098 (2023).
18. Przybył, K. *et al.* Application of machine learning to assess the quality of food products—case study: Coffee bean. *Appl. Sci.* **13**, 10786 (2023).

19. Ghabel, R., Rajabipour, A., Ghasemi-Varnamkhashti, M. & Oveisi, M. Modeling the mass of Iranian export onion (*Allium cepa* L.) varieties using some physical characteristics. *Res. Agric. Eng.* **56**, 33–40 (2010).
20. Babic, L., Matic-Kekic, S., Dedovic, N., Babic, M. & Pavkov, I. surface area and volume modeling of the williams pear (*Pyrus communis*). *Int. J. Food Prop.* **15**, 880–890 (2012).
21. Shahin, M. A. & Symons, S. J. A machine vision system for grading lentils. *Can. Biosyst. Eng.* **43**, 7 (2001).
22. Kılıç, K., Boyacı, I. H., Köksel, H. & Küsmenoğlu, İ. A classification system for beans using computer vision system and artificial neural networks. *J. Food Eng.* **78**, 897–904 (2007).
23. Guevara-Hernandez, F. & Gil, J. G. A machine vision system for classification of wheat and barley grain kernels. *Span. J. Agric. Res.* **3**, 672–680 (2011).
24. Tahir, A. R. *et al.* Evaluation of the effect of moisture content on cereal grains by digital image analysis. *Food Res. Int.* **40**(9), 1140–1145 (2007).
25. Wilson, J. D., Bechtel, D. B., Todd, T. C. & Seib, P. A. Measurement of wheat starch granule size distribution using image analysis and laser diffraction technology. *Cereal Chem.* **83**(3), 259–268 (2006).
26. Arasan, S., Akbulut, S. & Hasiloglu, A. S. Effect of particle size and shape on the grain-size distribution using Image analysis. *Int. J. Civ. Struct. Eng.* **1**(4), 968–985 (2011).
27. Horabik, J. Impact of humidity on physical and geometrical properties of slavic varieties of rye grains. *Acta Agroph.* **54**, 5–121 (2001) (in Polish).
28. Zewdu, A. D. & Solomon, W. K. Moisture-Dependent physical properties of Tef seed. *Biosyst. Eng.* **96**, 57–63 (2007).
29. Gierz, Ł., Kolankowska, E., Markowski, P. & Koszela, K. Measurements and analysis of the physical properties of cereal seeds depending on their moisture content to improve the accuracy of dem simulation. *Applied Sciences* **12**, 549 (2022).
30. Gupta, R. K. & Das, S. K. Physical properties of sunflower seeds. *J. Agric. Eng. Res.* **66**, 1–8 (1997).
31. Sacilik, K., Öztürk, R. & Keskin, R. Some physical properties of hemp seed. *Biosyst. Eng.* **86**, 191–198 (2003).
32. Altuntas, E. & Demirtola, H. Effect of moisture content on physical properties of some grain legume seeds. *NZ J. Crop Hortic. Sci.* **35**, 423–433 (2007).
33. Al-Kazaali, H. A. & Baktash, F. Y. Impact of corn grain moisture at harvesting to agronomic traits in subsequent generation. *Iraqi J. Agric. Sci.* **48**, 18 (2017).
34. Ekechukwu, O. V. Review of solar-energy drying systems I: An overview of drying principles and theory. *Energy Convers. Manag.* **40**(6), 593–613 (1999).
35. Omosekeji, G.M., 2018. Industrial vision robot with raspberry pi using pixy camera: stereo vision system.
36. Al-Sammarraie, M. A. J. & Özbek, O. Comparison of the effect using color sensor and pixy2 camera on the classification of pepper crop. *J. Mech. Eng. Res. Dev.* **44**, 396–403 (2021).
37. Nasution, T. H. *et al.* Analysis of the use of cmucam5 pixy camera in wheeled soccer robots. *IOP Conf. Ser. Mater. Sci. Eng.* **851**, 12034 (2020).
38. Assadzadeh, S., Walker, C. K. & Panozzo, J. F. Deep learning segmentation in bulk grain images for prediction of grain market quality. *Food Bioprocess Technol.* **15**, 1615–1628 (2022).
39. Gierz, Ł., Przybył, K., Koszela, K., Duda, A. & Ostrowicz, W. The use of image analysis to detect seed contamination—a case study of triticale. *Sensors* **21**, 151 (2020).
40. Kumar, M. N., Koushik, K. V. S. & Deepak, K. Prediction of heart diseases using data mining and machine learning algorithms and tools. *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.* **3**, 887–898 (2018).
41. Kasap, A. Physical properties of monogerm sugarbeet (*Beta vulgaris* var. altissima) seeds. *NZ J. Crop Hortic. Sci.* **34**, 311–318 (2006).
42. Garnayak, D. K., Pradhan, R. C., Naik, S. N. & Bhatnagar, N. Moisture-dependent physical properties of Jatropha seed (*Jatropha curcas* L.). *Industrial Crops Prod.* **27**, 123–129 (2008).
43. Tavakoli, M., Tavakoli, H., Rajabipour, A., Ahmadi, H. & Gharib-Zahedi, S. M. T. Moisture-dependent physical properties of barley grains. *Int. J. Agric. Biol. Eng.* **2**, 84–91 (2010).
44. Masyitah, M., Agussabti, A. & Kasimin, S. Tingkat adopsi petani terhadap benih unggul padi sawah di kabupaten aceh besar provinsi aceh. *Agrifo Jurnal Agribisnis Universitas Malikussaleh* **4**, 27–32 (2019).
45. Nashmi, R. J. & Naser, J. M. Improving the rheological and qualitative properties of bread wheat by barley-extracted pentosanes. *Iraqi J. Agric. Sci.* **53**, 1212–1222 (2022).
46. Kania, M. & Andrejko, D. Influence of different preparing methods of wheat grain before milling on flour moisture. *Acta Scientiarum Polonorum. Technica graria* **10**(1–2), 15–21 (2011).
47. Al-Sammarraie, M. A. J. *et al.* Predicting fruit's sweetness using artificial intelligence—case study: Orange. *Appl. Sci.* **12**, 8233 (2022).
48. Ramotra, A. K., Mahajan, A., Kumar, R. & Mansotra, V. Comparative analysis of data mining classification techniques for prediction of heart disease using the weka and spss modeler tools. In *Smart Trends in Computing and Communications* 89–97 (Springer, 2020).
49. Dwivedi, A. K. Performance evaluation of different machine learning techniques for prediction of heart disease. *Neural Comput. Appl.* **29**, 685–693 (2018).
50. Abrougui, K. *et al.* Modular feed forward networks to predict soil penetration resistance from tillage technique and working depth. *Int. J. Curr. Eng. Technol.* **4**, 3567–3573 (2014).
51. Al-Taie, A. J., Al-Bayati, A. F. & Taki, Z. N. M. Compression index and compression ratio prediction by artificial neural networks. *J. Eng.* **23**, 96–106 (2017).
52. Al-Sammarraie, M. A. J. & Kirilmaz, H. Technological advances in soil penetration resistance measurement and prediction algorithms. *Rev. Agric. Sci.* **11**, 93–105 (2023).
53. Al-Sammarraie, M. A. J., Al-Aani, F. & Al-Mashhadany, S. A. Determine, predict and map soil ph level by fiber optic sensor. *IOP Conf. Ser. Earth Environ. Sci.* **1225**(1), 012104 (2023).
54. Ibrahim, M. A., Mohammed-Ridha, M. J., Hussein, H. A. & Faisal, A. A. H. Artificial neural network modeling of the water quality index for the Euphrates river In Iraq. *Iraqi J. Agric. Sci.* **51**(6), 1572–1580 (2020).

Acknowledgements

This study was supported by the National Centre for Research and Development under the LIDER VIII programme, project No. LIDER/24/0137/L-8/16/NCBIR/2017. Publication co-financed by the state budget under the program of the Ministry of Education and Science (Republic of Poland) under the name Excellent Science—Support for Scientific Conferences entitled XXIII Polish Nationwide Scientific Conference “PROGRESS IN PRODUCTION ENGINEERING 2023” project number DNK/ SP/546290/2022 amount of funding PLN 162 650,00 total value of the project PLN 238 650,00 (Poland). The authors thank the breeders for providing the Main Seed Warehouse with PlantiCo Zielonki Sp. z o. o. Spójna, Breeding and Seed Plant in Nochowo, located in Greater Poland Province and allowing their plant material to be used for scientific and research purposes. on the legal protection of plant varieties of January 22, 2021 (Journal of Laws of 2021, item 213) belongs to him.

Author contributions

Ł.G., M.A.J.A and P.M. contributed equally to this work. Ł.G. and M.A.J.A.: conceptualization, data analysis and simulation; Ł.G., M.A.J.A, O.Ö. and P.M.: visualization; Ł.G. and P.M.: test and writing original manuscript; P.M.: statistical analysis of research results; Ł.G., O. Ö. and P.M.: basic review and editing; Ł.G.: resources; Ł.G., O.Ö. and P.M.: revise manuscript and check all data; Ł.G.: supervision; Ł.G.: project administration; Ł.G. and P.M.: funding acquisition.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Ł.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024