Research Article

Identification of Wheat Varieties with a Parallel-Plate Capacitance Sensor Using Fisher's Linear Discriminant Analysis

C. V. K. Kandala,[1] K. N. Govindarajan,[2] N. Puppala,[3] V. Settaluri,[4] and R. S. Reddy[4]

[1] *United States Department of Agriculture, Dawson, GA, USA*
[2] *Bio-Systems Engineering, University of Nebraska, Lincoln, NE, USA*
[3] *New Mexico State University, Clovis, NM, USA*
[4] *Department of Biotechnology, K.L. University, Vaddeswaram, India*

Correspondence should be addressed to C. V. K. Kandala; chari.kandala@ars.usda.gov

Fisher's linear discriminant (FLD) models for wheat variety classification were developed and validated. The inputs to the FLD models were the capacitance (C), impedance (Z), and phase angle (θ), measured at two frequencies. Classification of wheat varieties was obtained as output of the FLD models. Z and θ of a parallel-plate capacitance system, holding the wheat samples, were measured using an impedance meter, and the C value was computed. The best model developed classified the wheat varieties, with accuracy of 95.4%, over the six wheat varieties tested. This method is simple, rapid, and nondestructive and would be useful for the breeders and the peanut industry.

1. Introduction

Wheat (*Triticum aestivum* L.) is a prominent crop grown worldwide and also is one of the most important food items consumed in different forms such as bread, cookies, and pasta. It is also an important ingredient in hundreds of other food and drink preparations such as pizza, cakes, soups, and beer. Many wheat varieties exist, specific to the local cultures in different regions of the world. Wheat breeders developed hundreds of varieties to improve not only the yields, but also such agronomic and quality attributes such as resistance to pests and diseases and stability in height and growth. Over the past several years, producers in different geographical areas started preferring particular wheat varieties to produce end products such as bread or beer, according to the local tastes. Thus, variety identification plays an important role in selecting the right type of wheat for a particular product and assures its quality. Many product manufacturers such as bakeries and restaurants demand high levels of purity with respect to the variety. Techniques used presently for variety identification include gel electrophoresis and high performance liquid chromatography (HPLC). The CSIRO Plant Industry of Australia [1] developed a testing system using a set of DNA markers to identify wheat and barley varieties. However, these methods are time consuming and need some level of expertise to use. Thus, a physical method which is rapid and nondestructive would be useful for both the breeder and the industry, in maintaining the required quality of the wheat and its products.

2. Materials and Methods

It was found earlier that there exists a high correlation between dielectric properties of aqueous materials and their moisture content (MC). The variation in dielectric constant with MC for shelled yellow field corn was found to be more pronounced between 1 and 5 MHz [2]. Because the degree of change in the dielectric constant with the change in moisture content decreases with increasing frequency [3], the difference in the dielectric constants of a parallel-plate capacitor, holding wheat samples between the plates, measured at the two frequencies (1 MHz and 5 MHz) should be a good estimator of moisture content. Since capacitance is a function of dielectric constant, the capacitance difference at these two frequencies should also be a good indicator of

the moisture content [4]. However, attempts to estimate MC of wheat samples by a parallel-plate system as mentioned above, using the two frequencies, did not yield sufficiently accurate results [5]. This was partially because the volume of space that a sample of odd-shaped material, such as grain, occupies between two parallel plates would vary each time the material is placed between the plates. Air gaps between the grain kernels and between the kernels and the capacitor walls would occur differently, introducing errors. To compensate for these errors, two other related electrical parameters, phase angle (θ) and impedance (Z), were also measured at these two frequencies using the CI meter (CI meter is Chari's impedance meter designed and constructed by the corresponding author). The capacitance of the parallel-plate system was computed from the values of θ and Z. The dependence of MC on C, θ, and Z was earlier studied and found to be useful in MC determination for wheat [6]. Thus, the three parameters C, θ, and Z measured and/or computed at two frequencies could be used to determine the MC of six varieties of wheat with acceptable accuracy (within 1% of their air-oven values) [7]. Fisher's linear discriminant analysis (FLD) is a method of finding a coefficient, with which a linear combination of relevant variables can discriminate different groups [8]. FLD analysis was earlier used for beef tenderness classification [9]. In the present work, FLD models for wheat variety classification and identification were developed with the three variables C, θ, and Z that were earlier found to be useful for MC determinations. The models were tested and validated on six varieties of wheat.

2.1. The CI Meter.

The design and operation of the CI meter were described previously [10]. To describe briefly, three frequencies (1, 5, and 9 MHz) are generated by three FOX crystal oscillators (FOX Electronics, Fort Myers, FL, USA) made for these frequencies and applied alternately to a parallel-plate system holding the samples between the plates, which acts as the impedance load (Z), by switching through a multiplexer (Figure 1). These crystals (Model HC-49U) have a frequency stability of ±50 PPM, over their operating temperature range, and the circuits are similar for the three frequencies. Initially, at 1 MHz, the current flowing through the system with an impedance Z is fed into an op-amp.

The same current would flow through the feedback resistor R_r. The output voltage of the op-amp and the original signal from the oscillator are rectified and measured as e_{m1} and e_{r1}, respectively. The current through Z is calculated as e_{m1}/R_r and the magnitude of the impedance of the parallel-plate system with a sample between them is obtained as $|Z_1| = R_r(e_{r1}/e_{m1})$.

The phase angle at 1 MHz is determined by comparing the signal emerging out of the op-amp with that of the original signal, using a comparator and phase detector that give an output voltage e_{p1}, proportional to the phase angle θ_1 between the two. The computer then switches the multiplexer to 5 MHz, and impedance $|Z_2|$ and phase angle θ_2 are measured. The real and imaginary components of Z at each frequency are calculated as $R = |Z| \cos \theta$ and $X = |Z| \sin \theta$. The capacitance of the parallel-plate system with a sample
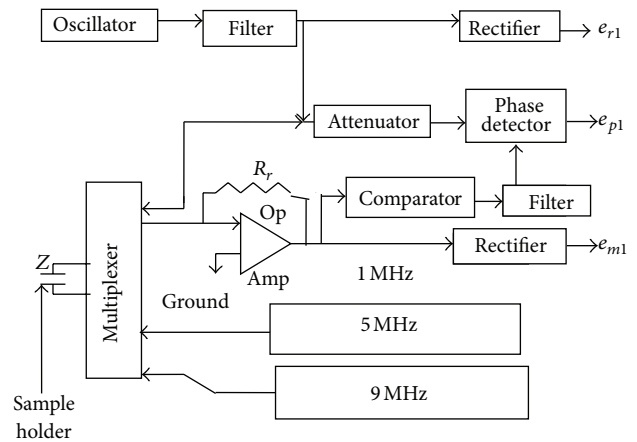

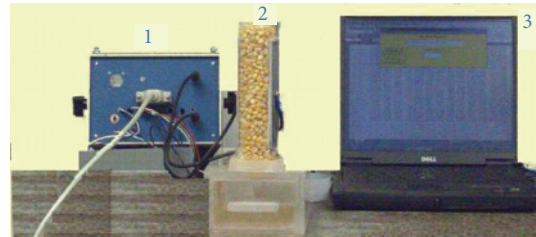Figure 1: RF impedance measuring block diagram.

Figure 2: RF impedance measuring system: (1) CI meter, (2) cylinder with wheat sample between electrodes, and (3) computer.

between the plates is obtained as $C = -1/2\pi fX$ at each frequency. The measurement system is shown in Figure 2. The CI meter is equipped with a regulated power supply that can be plugged into a 110 VAC line and two 12 VDC rechargeable batteries for field operations. A laptop computer controls the process and collects the data from the CI meter. The data is stored in the computer which was programmed to identify the wheat varieties. A cylindrical acrylic tube, fitted with a set of parallel-plate electrodes (Figure 2), served as the sample holder and sensor, as described earlier [10]. Inside the cylinder an electrode assembly consisting of two rectangular aluminum plates was fitted about 25 mm from the ends of the cylinder. The gap between the parallel-plates is filled with the sample, as shown in Figure 2. Except for the two electrodes, no metal parts were used in the assembly of the electrode system or in the sample collecting system, to prevent any interaction with the RF signal used in the measurements. With the drawer below the cylinder pushed all the way in, the cylinder was filled with the wheat sample, and the impedance measurements were taken. After completion of the measurements, the drawer was pulled out slowly, allowing the sample to fall into the drawer. The drawer was emptied before another sample was placed in the cylinder, for measurement. With a wheat sample occupying the space between the electrodes, the analyzer measured the impedance and phase angle of this electrode system at the three frequencies. The data was collected and subjected to FLD analysis.

2.2. Wheat Samples. Six varieties of wheat, planted and harvested around the Texas Panhandle and at the New Mexico State University station near Clovis, were used in this study [11]. The wheat varieties were Tam III, Duster, Scoutt 66, Endurance, Jagger, and Hatcher, planted during October 2010 and harvested during July 2011. All sample lots were stored at 4°C and 40% relative humidity. When the samples were received at the USDA-ARS National Peanut Research Laboratory (NPRL), their MC was about 9% (all moisture contents are expressed in percent wet basis in this paper). Each wheat variety was divided into two sublots one for training and the other for validation and was stored in separate airtight containers.

2.3. Procedures. Impedance measurements were made on 75 samples from each sublot of Duster, Endurance, Hatcher, and Jacqueline varieties, while measurements were made on 135 samples of Tam III and 60 samples of Scoutt 66 as per the availability of the samples. Each wheat sample was transferred from the container into the cylinder fitted with the electrode system until the space between the two parallel plates was completely filled. The cylinder accommodated about 150 g of wheat sample. The room temperature during the measurements was maintained at 21°C $\pm$ 1°C. With a sample in the cylinder, the impedance (Z) and phase angle (θ) were measured with the CI meter at 1, 5, and 9 MHz. The computer was programmed to repeat each measurement 30 times, compute the average value, and save it to an Excel spreadsheet. The sample was then collected in the drawer below the cylinder by gently pulling the drawer out and tapping on the cylinder for the sample to drop down. The drawer was emptied and reset in its box. This procedure was repeated for all wheat samples (in the two sublots) from the rest of the containers.

2.4. Data Analysis. Fisher's linear discriminant (FLD) models were developed using the statistical tool box of MATLAB (The Mathworks Inc., Natick, MA). The inputs to the FLD models are the capacitance (C), phase angle (θ), and the impedance (Z) measured for each of the wheat varieties with the impedance meter. The data set contained sample measurements of 495 each for training and validation, over the six varieties. To develop the FLD models, the 495 samples from the training sublot were used. The FLD models were tested with the 495 measurements made on the validation sublot. The overall classification accuracy obtained for the models in both training and validation processes was used as the evaluation metrics. Initially, models were developed using the C, θ, and Z values measured at 1 MHz and 5 MHz separately (the 9 MHz values were not used for wheat classification). Later, the measured values at the two frequencies were combined to obtain the best possible model.

3. Results and Discussion

Shown in Figure 3 are the classification results using the θ, Z, and C values at 1 MHz individually and with certain combinations as shown for both training and validation
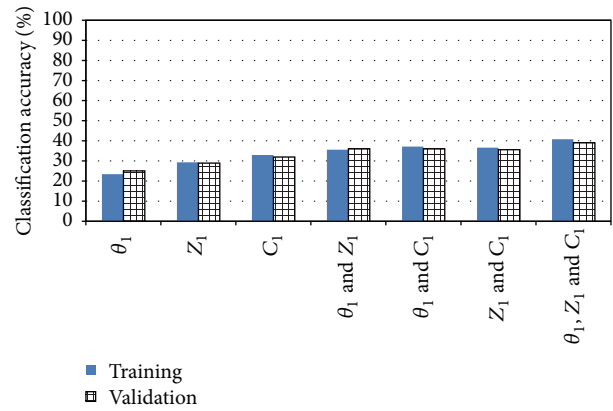

FIGURE 3: Models developed using different parameters at 1 MHz only.
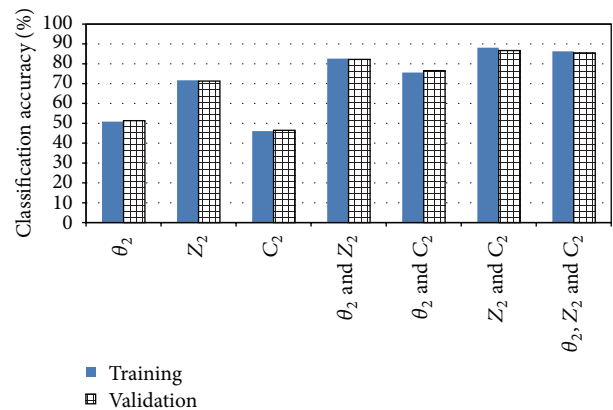

FIGURE 4: Models developed using different parameters at 5 MHz only.

groups. Though the classification accuracy improved with the three parameters combined, still it was not better than 40%. Shown in Figure 4 are the classification results obtained from the measurements at 5 MHz, using the same parameters individually and their different combinations as before. Though, the results showed considerable improvement at 5 MHz, the Z_2 and C_2 and θ_2, Z_2 and C_2 combinations giving an accuracy level of about 85%, it was felt that the combination of the measurements at the two frequencies would give even better classification accuracy than the parameters at only one frequency. Earlier, the combination of the measurements θ_1, Z_1 and C_1 and θ_2, Z_2 and C_2 at the two frequencies worked best for predicting MC of different varieties of wheat with a single calibration [7]. Thus, an attempt was made to perform the FLD analysis combining the θ, Z and C measurements at the two frequencies. Shown in Figure 5 are the classification results from the FLD analysis, combining both 1 MHz and the 5 MHz measurements of the three parameters. As expected the combination yielded better results. Both the training and validation sets showed a classification accuracy of 95% or better over the six wheat varieties. The classification accuracy of this model is shown in Table 1 for the training set and in Table 2 for the validation

TABLE 1: Classification performance of model using θ_1, Z_1, C_1, θ_2, Z_2, and C_2: training group.

Actual variety	Predicted variety						Total	Classification accuracy (CA%) within the group
	Duster	Endurance	Hatcher	Jacqueline	Tam III	Scoutt 66		
Duster	73	0	2	0	0	0	75	**97.3**
Endurance	1	74	0	0	0	0	75	**98.7**
Hatcher	2	2	71	0	0	0	75	**94.7**
Jacqueline	0	0	0	75	0	0	75	**100.0**
Tam III	15	0	0	0	120	0	135	**88.9**
Scoutt 66	0	5	0	0	0	55	60	**91.7**
Total	91	81	73	75	120	55	495	
CA (%) out of the 495 samples	**80.2**	**91.4**	**97.3**	**100.0**	**100.0**	**100.0**		**94.5**

TABLE 2: Classification performance of model using θ_1, Z_1, C_1, θ_2, Z_2, and C_2: validation group.

| Actual variety	Predicted variety						Total	Classification accuracy (CA%) within the group
	Duster	Endurance	Hatcher	Jacqueline	Tam III	Scoutt 66		
Duster	73	0	2	0	0	0	75	**97.3**
Endurance	0	75	0	0	0	0	75	**100.0**
Hatcher	1	0	74	0	0	0	75	**98.7**
Jacqueline	0	0	0	75	0	0	75	**100.0**
Tam III	15	0	0	0	120	0	135	**88.9**
Scoutt 66	0	5	0	0	0	55	60	**91.7**
Total	89	80	76	75	120	55	495	
CA (%) out of the 495 samples	**82.0**	**93.8**	**97.4**	**100.0**	**100.0**	**100.0**		**95.4**
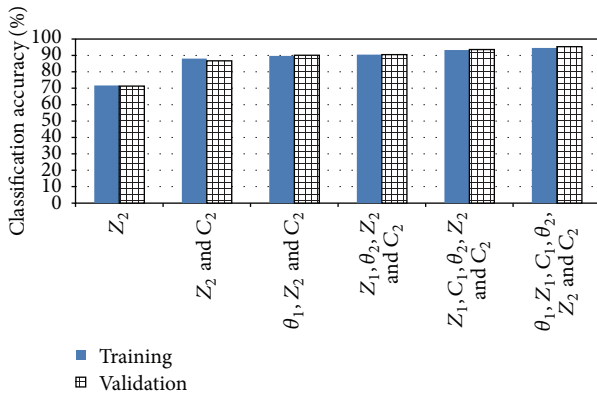
FIGURE 5: Models developed using different parameters at both 1 and 5 MHz.

set. The 495 samples used for validation were different from the ones used for training (calibration). The percentage of classification accuracy (CA%) shown in the last column in Tables 1 and 2 for the Duster variety, for example, is the number of samples that could be identified as Duster out of the total of 75 Duster samples, using the FLD analysis, in the training and the validation groups, respectively. For the Duster variety 73 samples out of 75 were identified as Duster, giving a classification accuracy of 97.3%. Similarly, the classification accuracy in the case of Jacqueline was 100% and was 89% for Tam III. For the other three varieties, the classification accuracy was over 90% as shown in the last column. The classification accuracy for the training group, over all the varieties, was 94.5%. The percentage of accuracy shown in the bottom row is the classification accuracy of each variety, predicted from all the varieties consisting of 495 samples. Thus, in the case of Duster a total of 91 samples were identified as Duster from the total of 495. The breakup was like 73 samples from Duster, 1 sample from Endurance, 2 from Hatcher, 15 from Tam III, and none from Jacqueline and Scott 66. This amounts to accuracy of 73 out of 91 identifications for Duster, which is 80.2%. Incidentally, Duster was the only variety with this low identification accuracy. All other varieties showed accuracy better than 91%. Similarly, from the validation sublot classification results, shown in Table 2, the classification percentage of accuracy shown in the last column for individual varieties, from the total number of samples of that variety predicted, stayed the same or slightly better, compared with the training group. The classification accuracy over all the varieties was over 95%.

Both Endurance and Jacqueline varieties showed a classification accuracy of 100% while the classification accuracy for Tam III remained the same at 89%. The other three groups showed accuracy of classification of over 90%. However, the classification accuracy numbers shown in the bottom row showed improvement over the training group classification accuracy. Duster variety showed 82% accuracy while all other varieties showed a classification accuracy of over 93% out of the 495 samples tested. The classification accuracy over all the varieties determined either within the group or out of the 495 samples (the last column or the bottom row in Tables 1 and

2) remains the same. Thus, the overall accuracy obtained in the classification of both the training and validation samples indicates the suitability of the FLD model, consisting of the capacitance, phase angle, and impedance measurements made at 1 and 5 MHz, as a useful tool in identifying the wheat varieties by this rapid and nondestructive method.

4. Conclusion

It is possible to identify several wheat varieties, from each other, using Fisher's linear discriminant (FLD) analysis. This method is rapid and nondestructive. Impedance measurements at two frequencies were made using a low-cost impedance analyzer designed and developed at the USDA laboratories by the corresponding author. There is a possibility of applying this method for identifying peanut varieties with high and low amounts of total oil or oleic acid contents. This type of identification would be useful to the breeder as well as the peanut industry.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] CSIRO, Enquiries, Bag 10, Clayton South, VIC, 3169, Australia.

[2] S. O. Nelson, "Frequency and moisture dependence of the dielectric properties of high-moisture corn," *Journal of Microwave Power*, vol. 13, no. 2, pp. 213–218, 1978.

[3] S. O. Nelson, "Frequency and moisture dependence of the dielectric properties of chopped pecans," *Transactions of the ASABE*, vol. 24, no. 6, pp. 1573–1576, 1981.

[4] S. Nelson, K. Lawrence, and C. Kandala, "Sensing moisture in peanut and pecan kernels by RF impedance measurements," *Journal of Microwave Power and Electromagnetic Energy*, vol. 27, no. 3, pp. 171–174, 1992.

[5] C. V. K. Kandala and S. O. Nelson, "Measurement of moisture content in single kernels of peanuts. A nondestructive electrical method," *Transactions of the ASAE*, vol. 33, no. 2, pp. 567–571, 1990.

[6] R. Thakur, S. Chatterji, A. Kumar, and B. S. Bansod, "Studies on electrical properties of wheat as a function of moisture content," *Quality Assurance and Safety of Crops and Foods*, vol. 3, pp. 198–204, 2011.

[7] C. V. K. Kandala, J. Sundaram, N. Puppala, and V. Settaluri, "Nondestructive measurement of moisture content of different varieties of wheat using a single calibration with a parallel-plate capacitor," *Transactions of the ASAE*, vol. 55, no. 4, pp. 1583–1587, 2012.

[8] R. Johnson and D. Wichern, *Applied Multivariate Statistical Analysis*, Pearson Prentice Hall, Upper Saddle River, NJ, USA, 6th edition, 2007.

[9] G. K. Naganathan, L. M. Grimes, J. Subbiah, C. R. Calkins, A. Samal, and G. E. Meyer, "Visible/near-infrared hyperspectral imaging for beef tenderness prediction," *Computers and Electronics in Agriculture*, vol. 64, no. 2, pp. 225–233, 2008.

[10] C. V. K. Kandala, C. L. Butts, and M. C. Lamb, "Moisture content determination for in-shell peanuts with a low-cost impedance analyzer and capacitor sensor," *Transactions of the ASABE*, vol. 51, no. 4, pp. 1377–1381, 2008.

[11] B. Bean, *Wheat Variety Trials Conducted in the Texas and New Mexico High Plains*, AgriLife Extension, College Station, Tex, USA, http://clovissc.nmsu.edu/documents/wheat-variety-trial-summary-final2011.