The *Journal of Cotton Science 24:112–120 (2020)*
http://journal.cotton.org, © The Cotton Foundation 2020

BREEDING AND GENETICS

Upland Cotton (*Gossypium hirsutum* L.) Fuzzy-Seed Counting by Image Analysis

Matthew T. Herritt, Don C. Jones, Alison L. Thompson*

ABSTRACT

In recent years cottonseed size has been reduced as a result of the substantial fiber yield increases cotton breeders have made. Small cottonseed size has been associated with reduced germination, low seedling vigor and stand establishment, and has created production problems for downstream whole seed users. The potential loss in revenue to the cotton industry due to small seed size is substantial and has prompted a renewed effort by breeders to generate high-yielding, high-quality varieties with increased seed size. To aid these efforts and enable a better understanding of the effects of seed characteristics on fiber, a fuzzy-seed imaging method was developed. The method utilizes inexpensive, off-the-shelf equipment and an open source image processing pipeline to derive the number of seeds, seed index, and seed area, height, width, and perimeter. The time to image the seed and process the image takes less than three minutes per sample on average. The seed counts and seed index were strongly correlated with manual measurements at *r* = 0.967 and 0.693, respectively. Associations among seed characteristics and fiber indicate seed area, when used to calculate lint density, could be a useful selection criterion for breeders to increase both yield and seed size.

Upland cotton (*Gossypium hirsutum* L.) is a critical natural fiber source that supports a multibillion-dollar (USD) global industry. The United States is the third largest producer of cotton fiber worldwide (Cotton Incorporated, 2020). Most U.S.-produced cotton is exported to foreign markets to produce woven fabrics and yarn for apparel and home goods (National Cotton Council, 2020). Over the years, the demand for high-quality fiber has increased to meet global textile manufacturing needs (Smith and Coyle, 1997). In response to these demands, cotton breeding programs have focused on increased yields while increasing the quality of fiber properties (Bridge et al., 1971; Campbell et al., 2011).

Although breeders have made tremendous gains in lint yield, cottonseed size and weight has decreased (Campbell et al., 2011; Culp and Harrell, 1975). The reduction in cottonseed size has caused concern within the cotton community as smaller seeds tend to have lower germination rates and seedling vigor, which relate to poor stand quality and lint yield (Snider et al., 2016). Reduced seed size has also caused problems at gins by contaminating the ginned fiber and at oil mills by reducing the ability to dehull the seed (Dowd et al., 2018). Overall, the potential for loss in revenue associated with reduced seed size is substantial. Recent studies have found that larger seed size can overcome low germination rates and is associated with increased seed oil content (Edmisten, 2015; Hinze et al., 2015). As a result, the National Cotton Council has asked the breeding community to consider seed size, particularly seed index, when developing new germplasm.

One of the primary reasons attributed to lint yield gains and reduced seed size is the heavy reliance on lint percentage as a selection criterion in breeding programs (Bridge et al., 1971; Groves and Bourland, 2010). The negative association between seed size and lint percentage is attributed to nutrient competition between the two, where breeders have driven partitioning towards developing fibers rather than developing seed (Campbell et al., 2011; Kloth and Turley, 2010). Prior to the extensive use of lint percentage, breeders relied on lint density, the amount of fiber on a unit area of the seed (Breaux, 1954), as a selection criterion. Cook (1908) hypothesized that use of lint density as a selection criterion could improve yield stability by standardizing seed size, and that larger seeds would have the most lint because larger seeds have more lint-bearing surface potential. However, Groves and Bourland (2010) pointed out

M.T. Herritt and A.L. Thompson*, U.S. Department of Agriculture, Agricultural Research Service, U.S. Arid-Land Agricultural Research Center, Maricopa, AZ 85138; and D.C. Jones, Cotton Incorporated, Cary NC 27513
*Corresponding author: alison.thompson@usda.gov

that an easy and accurate estimate of seed surface area was required to make lint density a feasible selection criterion for breeding programs. They developed a viable method using the fuzzy-seed index and an applied regression equation to estimate seed surface area (Groves and Bourland, 2010).

To improve cotton fiber quality, yield, and seed size, a quick and easy method to quantify seed characteristics is needed. To minimize time and cost, the method needs to be amenable to fuzzy seed. The primary goal of the present study was to develop a quick, easy, and accurate method to assess cottonseed characteristics. The objectives were to 1) utilize inexpensive, off-the-shelf imaging equipment to develop a fuzzy-seed imaging method; 2) compare traditional seed estimates with imaged seed outputs; and 3) associate imaged seed characteristics with fiber quality, yield, and seed germination.

MATERIALS AND METHODS

Cotton Germplasm and Field Experimental Design. Twenty-four upland cotton (*G. hirsutum*) elite breeding germplasm lines provided by the Regional Breeders Testing Network (RBTN) (Table 1) were planted at the University of Arizona, Maricopa Agriculture Center (33.068° N, 11.971° W, 360 m above sea level) in Maricopa, AZ on 20 May 2019. The experimental design was a randomized, complete block design with four replicates for each cotton entry. Experimental plots included two 10.6-m cotton rows with 1.02-m inter-row spacing and a density of approximately 8.6 plants m^2. The entries were furrow flood irrigated to germinate seed and establish plants, then switched to micro-irrigation via buried drip tape (Netafim, Fresno, CA) on 6 June 2019. Irrigations were scheduled from a daily crop-water use and soil-water balance model based on Food and Agriculture Organization-56 (FAO-56) methods (Allen et al., 1998; Hunsaker et al., 2005). Entries received 100% of the recommended irrigation amounts from 6 June until 20 September 2019 when irrigation was terminated. Liquid ammonium nitrate (UAN 32-0-0) was uniformly applied in three split applications on 15 May (pre-plant), 26 June, and 9 July 2019, totaling 150 kg N ha^{-1}. Experimental plots were defoliated on 14 October 2019, then harvested with a two-row cotton picker and plot weights recorded on 18 November 2019.

Table 1. The Regional Breeders Testing Network germplasm identified by the entry ID and lists the last name of the breeder cooperator and location of breeding program

Entry ID	Cooperator	Location
LA16063019	Myers	Alexandria, LA
LA16063033	Myers	Alexandria, LA
LA16063054	Myers	Alexandria, LA
13AFX6-27-2	Hinze	College Station, TX
13AFX13-12-5	Hinze	College Station, TX
Ark 1115-36	Bourland	Keiser, AR
Ark 1102-55	Bourland	Keiser, AR
Ark 1114-21	Bourland	Keiser, AR
Ark 1117-60	Bourland	Keiser, AR
Ark 1124-50	Bourland	Keiser, AR
Ark 1112-59	Bourland	Keiser, AR
TAM 13S-03	Smith	College Station, TX
TAM 12J-39	Smith	College Station, TX
TAMLBB15905	Dever	Lubbock, TX
TAMLBB16507	Dever	Lubbock, TX
GA2016024	Chee	Tifton, GA
GA2016099	Chee	Tifton, GA
GA2016103	Chee	Tifton, GA
MS 2010-87-37	Wallace	Mississippi State, MS
CSX8308	Jones	CSIRO, Australia
DP 393 CK	Wallace	Mississippi State, MS
DP 493 CK	Wallace	Mississippi State, MS
FM 958 CK	Wallace	Mississippi State, MS
UA 222 CK	Wallace	Mississippi State, MS

Fiber and Lint Analysis. Prior to mechanical harvest, two sets of 25 bolls were handpicked from each experimental plot on 4 November 2019. The 25-boll samples were weighed, then lint was separated from seed with a Table-Top 10-saw gin (Compass Systems, Barberton, OH). The lint weight and seed weight were recorded. Lint percentage turnout, traditional seed index (TSI), seed per boll, lint index, and cotton fiber yield were calculated as follows:

$$\text{Lint percent} = \frac{\text{lint weight (g)}}{\text{seed cotton weight (g)}} \text{x100}$$

$$\text{TSI} = \text{25 fuzzy seed weight (g) x 4}$$

$$\text{Seed per boll} = \frac{\text{seed weight (g)} / \left(\text{25 fuzzy weight (g)} / 25\right)}{25}$$

$$\text{Lint index} = \frac{\text{lint weight (g)}}{\text{seed per boll x25}}$$

$$\text{Lint yield} = \frac{\text{plot weight (kg) x lint percent x plot area (ha)}}{100}$$

A 10-g subsample from the first 25-boll sample from each plot was sent to Cotton Incorporated (Cary, NC) for fiber quality assessment on the HVI (USTER® AFIS PRO, Charlotte, NC) including fiber uniformity (%), elongation (%), micronaire (unit), strength (kN m kg^{-1}), and length (upper-half mean, mm). A 50-g subsample from the second 25-boll sample from each plot was sent to Texas Tech University (Lubbock, TX) for fiber quality assessment on the HVI. Both sets of HVI data were used for analysis. Fuzzy seed from both ginned 25-boll sample sets were retained in a paper bag at ambient temperature (24-26 °C) until image analysis.

Seed Image Analysis. To prepare for imaging, a Bounce® dryer sheet (P&G Profession, Cincinnati, OH) was placed in the paper bag with the fuzzy seed and baked in an oven at 55 °C for 48 h. Approximately 5 g of the baked fuzzy seed was taken from each bag and the weight recorded before spreading the seeds onto a custom-made tray (Fig. 1). Fuzzy seeds were spread to ensure that no seeds were touching, then scanned with a ScanSnap SV600 overhead scanner (FUJITSU, Tokyo, Japan). Note, the dryer sheet step is not necessary for imaging, but it does enable quicker seed spreading by removing static from the linters. The seeds on the tray were demarcated into two groups of 10 seeds for germination assays under warm and cool conditions. The demarcated seeds were transferred to germination trays lined with paper towels. At the start of the assay, the paper towels were moistened and another set of damp towels were placed over the fuzzy seed. The trays were then transferred to a germination cabinet (Percival Scientific, Perry, IA) and checked every day for 10 d to record germination. The warm germination assays were conducted at 26 °C daytime and 21 °C nighttime temperatures with 14-h daylengths. The cool assays were conducted at 25 °C daytime and 18 °C nighttime temperatures with 14-h daylengths. The percentage germination was calculated after 10 d. Germination efficiency was calculated as a weighted summation where seeds that germinated on Day 1 received a higher weight than those that germinated on Day 10. The average percent germination and efficiency was calculated for each RBTN entry.

$$\text{Efficiency} = \#\text{seed}\,(0.1) + \#\text{seed}\,(0.09) + \#\text{seed}(0.08) \ldots \#\text{seed}\,(0.03) + \#\text{seed}(0.02) + \#\text{seed}\,(0.01)$$

Figure 1. Custom built tray used for fuzzy cottonseed imaging and the ScanSnap overhead scanner.

Images of the fuzzy seeds were analyzed using a custom macro written for ImageJ (imagej.nih.gov) using ImageJ Macro Language (IJM). The macro searches for files with .jpg extensions in subfolders contained within a parent folder selected by the user. When a file with the .jpg extension is identified, the RGB (red, green, blue) image of the fuzzy seed is split into hue, saturation, and intensity layers. The getHistogram and setThreshold statements are used to apply a dynamic threshold to the saturation layer that identifies the lower 5% of pixels in the image. From the pixels identified by the dynamic threshold, seeds are identified with the Analyze Particles function of ImageJ. The size and circularity range used to identify seed is 0.2 to 3.0 and 0.25 to 1.00, respectively (Fig. 2). The size range for the algorithm was determined by measuring more than 5,000 seeds by hand with a micrometer. The circularity range for the algorithm was determined by performing several tests within ImageJ on fuzzy seed until only seeds were identified as "particles". The ImageJ software uses a set scale (1 pixel = 0.012 cm) to convert the number of pixels in each identified particle to calculate the seed parameters. ImageJ records the parameters of each particle identified (fuzzy seed), and the macro saves this output and moves on to the next file identified with the .jpg extension. The output file contains the number of seeds on the tray and the area, height, width, and perimeter for each seed. An imaged seed index (ISI) was calculated as:

$$\text{ISI} = \frac{\text{sample weight (g)}}{\text{number of seed}} \text{ x100}$$

$$\text{Lint density} = \frac{\text{lint index}}{\text{imaged seed area}}$$

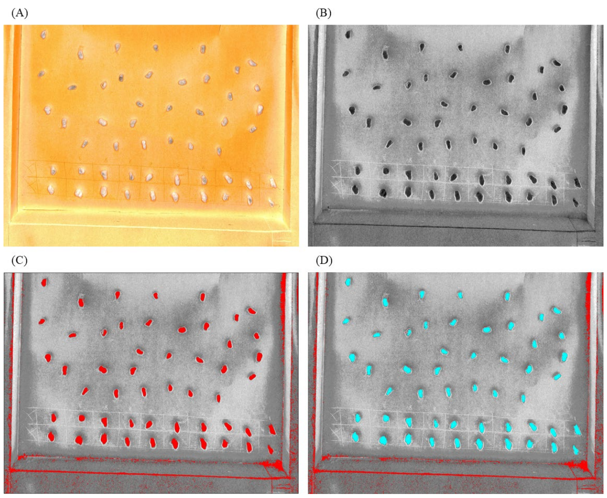

Figure 2. The image analysis workflow where (A) shows the initial red|green|blue image, (B) shows the saturation layer of the image, (C) shows the dynamic pixel threshold, and (D) shows the particle identification.

Seed Oil and Protein Content. RBTN serves as a multi-environment trial funded by Cotton Incorporated's Variety Improvement project (rbtn.cottoninc.com). In 2019, 16 locations across the cotton belt participated. Of these locations, seed from Florence, SC; Lubbock, TX; Mississippi State, MS; and Tallassee, AL were sent to the University of North Texas (Denton, TX) for seed oil and protein analysis. Each seed sample was measured in triplicate using the time domain 1H-NMR method developed by Horn et al. (2011). The average seed and oil protein content by weight percentage across locations was provided for each RBTN entry.

Statistical Analysis of Lint and Seed Traits. Lint yield, fiber quality traits, and germination were regressed against TSI, ISI, lint density, and the imaged seed characteristics. Pearson correlation coefficients and associated *p*-values were calculated using the stat.linregress function in the sci.py package for Python 3.0. Regression analysis did not include the seed oil and protein content as the data were not collected from the same seed. The matplotlib.pyplot package was used to generate box plots and identify outliers using the default fliers function, where points greater than or less than the first and third quartile were considered an outlier. Because more rigorous testing was not performed to identify outliers, those data were retained for the regression analysis and summary statistics. The summary statistics (mean, standard deviation from the mean, minimum value, and maximum value) were calculated using the describe function in the Pandas package for Python 3.0.

RESULTS AND DISCUSSION

Imaged Seed Assessment. The preparation and scan time for each seed tray was approximately 45 sec. Seeds not baked with the dryer sheet took approximately 150 sec to prepare and scan the seed tray. The ImageJ analysis took 120 sec per image. Utilizing fuzzy seed prevents some of the concerns regarding seed sample integrity, including seed fragments and loss of seed when using delinted seed. The ImageJ seed count output, when compared to the hand count, was highly accurate ($r = 0.967$, p-value $= 0.000$). Although a well-trained individual could likely count the number of seeds in the 5-g subsample faster than the time it took to spread, scan, and analyze the image, human error can be introduced and a bias towards small or immature seeds can be introduced. The advantage with the imaging method is that retained images can be rechecked for accuracy and bias can be reduced or eliminated. It is also possible to scan larger seed sample sizes, which would be faster than an individual counting and improve within-sample variation estimates. The efficiency of this method could also be improved by developing more trays, so that one person can be preparing the trays while another scans. Improving the ImageJ analysis time could be achieved on a computer with more processing units.

The ISI showed a reduced range in values (7.32-12.02) compared to TSI (5.92-13.20) and fewer outliers were identified in the ISI than the TSI, indicating the ISI method is more consistent for calculating seed index (Fig. 3). The decreased variation with the imaging method could be from one of two sources. First, the imaging method relies on a larger subsample, approximately 50 to 60 seeds, compared to the 25 seeds used in the traditional method, which enables a better representation of true variation. Second, any error in counting the 25 fuzzy seeds will become compounded when multiplying by 4 to get the 100 fuzzy-seed weight. As stated above, the chances of this error occurring with the imaging method is greatly reduced as users can go back and check any values that stand out.

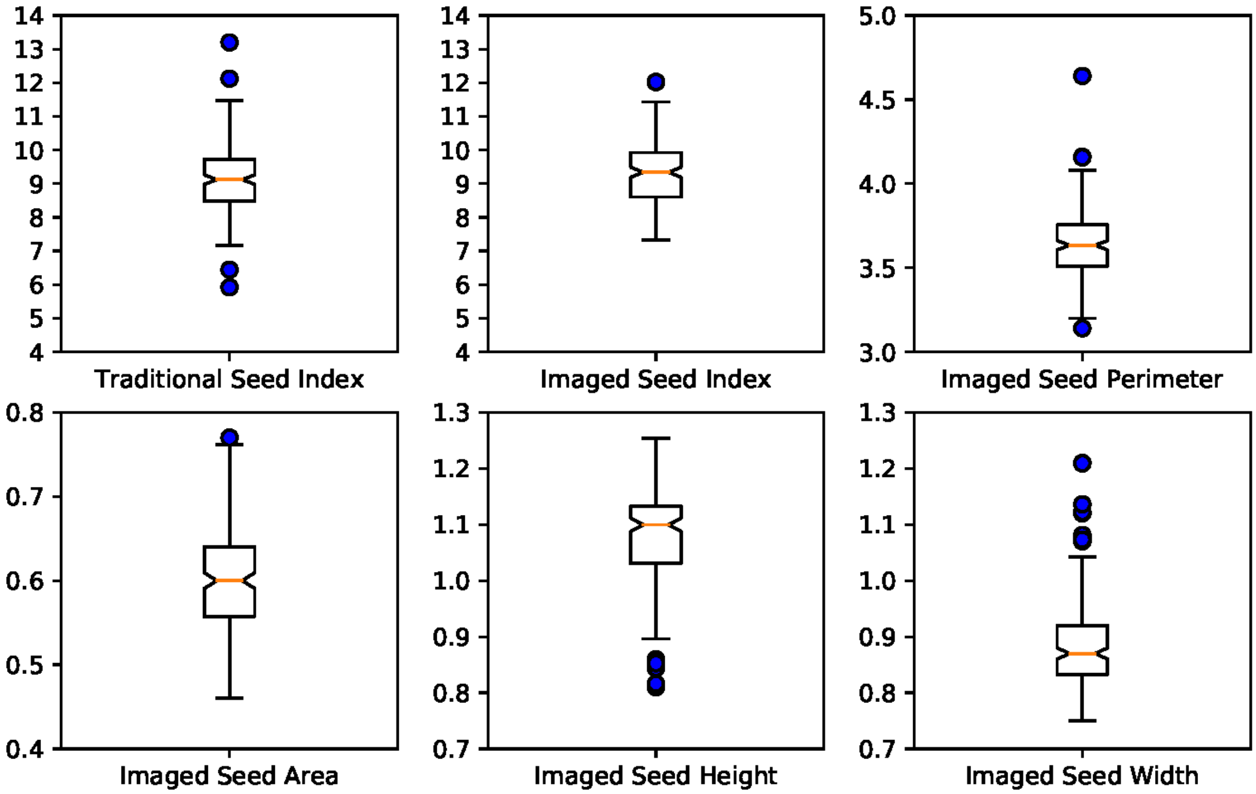

Figure 3. The boxplots showing the mean, quartiles, standard deviation from the mean, and outliers (blue) for the measured seed characteristics.

The seed perimeter, area, height, and width all showed variation within the sample seed set as expected (Fig. 3). Because the seeds were not specifically oriented on the tray, the height and width parameters were interchangeable so the difference between average seed height (1.08 cm) and width (0.88 cm) is relatively small and outliers were identified for both traits (Fig. 3). The boxplot analysis identified one outlier for the imaged seed area, which ranged from 0.46 to 0.72 cm^2 (Fig. 3). The imaged traits (seed index, area, perimeter, width, and height) were all positively and significantly associated with TSI (Table 2). Manual measurements of these parameters take considerable time and several methods have been developed, particularly to estimate seed surface area (Groves and Bourland, 2010; Hodson, 1920). The value added by estimating seed height, width, perimeter, and area with this method is tremendous as it provides breeders opportunities to better understand the partitioning of resources to cottonseed and fiber and the effects on fiber quality, fiber yield, and seed germination.

Association of Seed Characteristics with Fiber Quality and Yield. The TSI and ISI showed similar associations among the fiber quality traits where TSI showed stronger associations with micronaire and ISI with fiber strength and both showed a significant negative association with short fiber content (Table 3). These relationships are similar to previous findings and indicate the imaged seed index is a viable alternative for breeders (Desalegn et al., 2009; Kothari et al., 2015). The seed area showed similar associations with fiber quality traits as both ISI and TSI except for short fiber content. Lint density showed a significant positive association with micronaire and significant negative associations with length and strength that are consistent with findings by Smith and Coyle (1997). The seed perimeter showed significant associations with uniformity and strength, whereas seed width showed no other significant associations, indicating that seed area is more useful than the other imaged characteristics to understand seed and fiber quality dynamics.

Table 2. Pearson correlation coefficients and associated significance values for the measured characteristics

	TSI	ISI	Area	Perimeter	Width	Height
TSI	1.000					
p-value	0.000					
ISI	0.693	1.000				
p-value	0.000	0.000				
Area	0.555	0.690	1.000			
p-value	0.000	0.000	0.000			
Perimeter	0.509	0.763	0.817	1.000		
p-value	0.000	0.000	0.000	0.000		
Width	0.313	0.421	0.604	0.526	1.000	
p-value	0.000	0.000	0.000	0.000	0.000	
Height	0.346	0.464	0.580	0.549	-0.269	1.000
p-value	0.000	0.000	0.000	0.000	0.000	0.000

Table 3. Pearson correlation coefficients and associated significance values for the measured seed characteristics and the fiber yield, quality, and seed germination traits

Traditional Seed Index												
Trait	**Yield**	**LPCT**	**MIC**	**UHM**	**UI**	**STR**	**ELO**	**SFC**	**% Germ W**	**Efficiency W**	**% Germ C**	**Efficiency C**
***r* value**	-0.096	-0.596	-0.316	0.220	0.204	0.283	0.107	-0.250	-0.011	-0.067	0.084	-0.037
***p* value**	0.187	0.000	0.000	0.003	0.005	0.000	0.146	0.014	0.900	0.430	0.319	0.660
Imaged Seed Index												
Trait	**Yield**	**LPCT**	**MIC**	**UHM**	**UI**	**STR**	**ELO**	**SFC**	**% Germ W**	**Efficiency W**	**% Germ C**	**Efficiency C**
***r* value**	-0.115	-0.507	-0.155	0.227	0.275	0.344	0.090	-0.353	-0.005	-0.071	0.132	0.042
***p* value**	0.112	0.000	0.035	0.002	0.000	0.000	0.224	0.000	0.950	0.404	0.116	0.620
Imaged Seed Area (cm^2)												
Trait	**Yield**	**LPCT**	**MIC**	**UHM**	**UI**	**STR**	**ELO**	**SFC**	**% Germ W**	**Efficiency W**	**% Germ C**	**Efficiency C**
***r* value**	-0.269	-0.575	-0.301	0.240	0.098	0.357	-0.020	-0.132	0.038	-0.180	0.185	-0.024
***p* value**	0.000	0.000	0.000	0.001	0.182	0.000	0.791	0.201	0.658	0.033	0.027	0.776
Lint Density (g/cm^2)												
Trait	**Yield**	**LPCT**	**MIC**	**UHM**	**UI**	**STR**	**ELO**	**SFC**	**% Germ W**	**Efficiency W**	**% Germ C**	**Efficiency C**
***r* value**	0.273	0.600	0.310	-0.199	0.032	-0.216	0.121	-0.062	0.014	0.138	-0.128	-0.022
***p* value**	0.000	0.000	0.000	0.006	0.662	0.003	0.100	0.552	0.868	0.104	0.129	0.797
Imaged Seed Perimeter (cm)												
Trait	**Yield**	**LPCT**	**MIC**	**UHM**	**UI**	**STR**	**ELO**	**SFC**	**% Germ W**	**Efficiency W**	**% Germ C**	**Efficiency C**
***r* value**	-0.098	-0.375	-0.083	0.101	0.180	0.272	-0.042	-0.186	-0.032	-0.187	0.133	-0.034
***p* value**	0.178	0.000	0.260	0.169	0.014	0.000	0.568	0.070	0.704	0.027	0.114	0.691
Imaged Seed Width (cm)												
Trait	**Yield**	**LPCT**	**MIC**	**UHM**	**UI**	**STR**	**ELO**	**SFC**	**% Germ W**	**Efficiency W**	**% Germ C**	**Efficiency C**
***r* value**	-0.050	-0.230	-0.055	0.039	-0.040	0.084	-0.055	-0.047	0.093	-0.159	0.147	0.021
***p* value**	0.491	0.001	0.453	0.599	0.592	0.255	0.454	0.650	0.277	0.061	0.081	0.807
Imaged Seed Height (cm)												
Trait	**Yield**	**LPCT**	**MIC**	**UHM**	**UI**	**STR**	**ELO**	**SFC**	**% Germ W**	**Efficiency W**	**% Germ C**	**Efficiency C**
***r* value**	-0.228	-0.412	-0.267	0.231	0.200	0.317	0.044	-0.159	-0.094	-0.059	0.057	-0.041
***p* value**	0.001	0.000	0.000	0.002	0.006	0.000	0.552	0.122	0.268	0.489	0.504	0.629

TSI, ISI, and other seed traits, except for lint density, were negatively associated with lint percentage, which is consistent with findings by Hodson (1920), Breaux (1954), Bridge et al. (1971), Culp and Harrell (1975), and Campbell et al. (2011). Both seed indices and seed area had significant negative associations with fiber yield (Table 3). The negative association between seed area and yield was expected as the seed area was positively associated with seed index and is consistent with findings by Breaux (1954); however, seed area could have the potential to increase yield. As Cook (1908) explained, increasing seed surface area increases the lint-bearing surface potential. The calculated lint density (lint index/seed area) in this study showed a significant positive association with fiber yield at $r = 0.273$, p-value = 0.000, whereas the association between seed index and fiber yield was $r = 0.212$, p-value = 0.003. A much larger study is needed to determine the validity of this association and how useful it will be for a breeding program.

Association of Seed Characteristics with Germination. No significant associations were found between the percentage germination or germination efficiency under warm conditions with the TSI or ISI (Table 3). Only the imaged seed area had a significant positive association with percentage germination under cool conditions. The seed area and perimeter were the only two imaged seed characteristics found to have significant negative associations with germination efficiency under warm conditions, in other words, larger seeds by area were slower to germinate (Table 3). In cotton, larger seeds by mass (seed/kg) have been shown to germinate more quickly and have increased seedling vigor (Pettigrew and Meredith, 2009; Snider et al., 2016). However, Krieg and Bartee (1975) found that seeds with a higher density (g/cm^3) were restricted in early water imbibition, which resulted in reduced initiation rates of germination but not percentage germination, even at warmer temperatures. As seeds that had a larger area also had higher weights, it is likely these seeds also have increased density, so the negative association between germination efficiency and seed area could be due to slow imbibition. It will be important for breeders to note the distinctions in seed size characteristics and germination rates to ensure larger seeds with strong germination potential are carried forward in their programs.

Large and Small Seed Lines Identified by Image Analysis. ISI and TSI both identified DP 493 CK as having the smallest seeds of the entries tested at 7.73 and 7.25, respectively, whereas Ark 1115-36 had the smallest seeds based on seed area (0.49 cm^2). TSI identified 13AFX6-27-2 as the entry with the largest seeds at 10.70, whereas ISI identified TAM 12J-39 as the largest at 11.05. The seed area identified TAMLBB16507 and 13AFX6-27-2 with the largest seeds at 0.72 cm^2 (Table 4). In general, entries with increased seed oil content (> 19.5 %w) had increased TSI, ISI, and area values, whereas entries with increased seed protein content (> 19.5 %w) only had increased TSI or ISI values (Table 4). These findings are similar to Hinze et al. (2015) who reported a significant positive association between seed index and seed oil content in an upland cotton diversity panel. However, Hinze et al. (2015) also reported a significant negative association between seed index and seed protein content. The relationship found in this study would suggest cotton breeders have made significant improvements for seed protein content in modern germplasm.

Lint density identified Ark 1115-36 with the most fiber-per-seed area (> 0.130 g/cm^2), whereas TAMLBB16507 and 13AFX6-27-2 had the least (< 0.090 g/cm^2). The top five highest yielding germplasm included Ark 1115-36, whereas TAMLBB16507 was in the bottom five (Table 4). If the hypothesis made by Cook (1908) were correct, the yield of Ark 1115-36 could increase with larger seed size, whereas TAMLBB16507 has the potential for higher yields based on seed area. Further studies are needed to identify the heritability and environmental interactions for lint density that will aid a breeder's decision on the usefulness of lint density as a selection criterion. The development of this simple, efficient imaging method and analysis pipeline will enable such studies to be conducted.

CONCLUSION

This study describes an inexpensive, quick, and accurate method to quantify cottonseed traits from fuzzy seed for use in breeding programs. Analysis using the fuzzy-seed imaging method identified less variation and fewer outliers in the cotton imaged seed index compared to the traditional seed index, which indicates the imaging method is less prone to human error. The imaging method also provides opportunities to calculate seed characteristics that were previously too time intensive to consider in large breeding programs.

Table 4. The calculated mean for each germplasm entry across replicates for each of the measured traits

Entry	Weight (g)	Number of Seed	Area (cm2)	Perimeter (cm)	Width (cm)	Height (cm)	Imaged seed index	Traditional seed index	Lint density (g/cm2)	Lint Pct (%)	Yield (Kg/ ha1)	Micronaire	Length (mm)	Uniformity (%)	Strength (kN m kg ha1)	Elongation (%)	Shor fiber content (%)	Germination W (%)	Efficiency W	Germination C (%)	Efficiency C	[z]Seed Oil (%w)	[z]Seed Protein (%w)
13AFX13-12-5	5.01	55.50	0.59	3.64	0.87	1.09	9.05	9.54	0.095	36.54	861.08	4.65	1.12	81.14	30.39	5.23	9.80	76.67	0.96	73.33	0.81	19.26	20.10
13AFX6-27-2	5.04	47.75	0.72	3.98	1.00	1.14	10.56	10.70	0.089	36.78	808.26	5.10	1.20	83.21	37.10	6.23	8.13	86.67	1.06	81.67	0.69	19.88	19.89
Ark 1102-55	5.01	50.88	0.56	3.59	0.91	1.01	9.87	9.81	0.122	40.14	826.68	4.90	1.12	81.45	27.74	5.69	9.55	85.00	1.19	76.67	1.17	21.36	21.75
Ark 1112-59	5.02	50.00	0.64	3.88	0.92	1.13	10.06	9.56	0.109	41.42	865.57	5.12	1.10	82.13	30.23	5.66	8.95	85.00	0.98	68.33	0.65	19.50	20.29
Ark 1114-21	5.06	61.00	0.57	3.55	0.86	1.05	8.32	8.30	0.108	41.77	805.30	4.73	1.11	81.20	29.05	5.68	9.30	98.00	1.60	81.67	0.98	17.20	19.29
Ark 1115-36	5.03	58.88	0.49	3.39	0.80	1.00	8.55	7.85	0.136	44.93	942.24	5.37	1.13	83.39	30.37	6.23	8.18	90.00	1.67	63.33	1.14	17.09	19.88
Ark 1117-60	5.05	54.75	0.58	3.59	0.90	1.04	9.34	9.39	0.117	41.35	907.55	4.99	1.14	82.50	31.46	5.74	8.73	90.00	1.23	58.57	0.52	19.36	19.34
Ark 1124-50	5.04	54.00	0.61	3.71	0.91	1.08	9.33	8.82	0.109	42.23	743.62	5.12	1.14	82.51	30.36	5.41	8.78	76.67	0.88	85.00	0.85	16.96	18.89
CSX8308	5.02	62.88	0.53	3.47	0.86	1.02	8.00	7.93	0.126	45.23	855.34	5.18	1.15	80.86	31.58	5.35	9.38	78.33	0.94	78.33	0.71	16.47	18.43
DP 393 CK	5.04	54.38	0.57	3.62	0.85	1.09	9.30	8.87	0.125	41.66	900.49	5.11	1.10	82.04	30.85	6.14	8.68	91.67	0.99	76.67	0.75	18.17	19.86
DP 493 CK	5.04	65.63	0.50	3.36	0.83	0.97	7.73	7.25	0.119	44.08	981.64	5.58	1.04	79.86	27.36	5.31	10.45	84.00	0.87	46.67	0.31	17.99	19.63
FM 958 CK	5.05	53.00	0.64	3.73	0.91	1.12	9.56	9.70	0.098	38.52	683.36	4.98	1.10	81.85	28.96	5.26	9.00	56.67	0.38	60.00	0.54	19.83	20.44
GA2016024	5.06	53.25	0.62	3.77	0.87	1.14	9.53	9.64	0.112	41.23	667.70	4.79	1.12	82.04	31.11	5.46	8.63	91.67	1.10	75.00	0.70	17.02	19.61
GA2016099	5.04	56.13	0.61	3.59	0.92	1.05	8.99	9.04	0.109	41.59	822.56	4.97	1.13	82.01	32.17	6.01	8.98	94.29	1.20	57.14	0.66	16.65	18.34
GA2016103	5.01	63.63	0.53	3.32	0.81	1.02	7.89	8.25	0.115	41.30	814.23	5.40	1.11	81.41	30.92	5.64	9.50	91.43	1.24	57.14	0.54	17.39	18.95
LA16063019	5.04	57.25	0.59	3.60	0.87	1.08	8.81	8.51	0.111	43.04	173.43	5.18	1.14	82.03	33.60	5.80	8.88	71.67	0.85	76.67	0.62	na	na
LA16063033	5.01	53.88	0.62	3.67	0.86	1.13	9.33	9.17	0.100	39.20	373.72	4.98	1.14	83.01	33.09	5.93	8.43	84.00	0.79	72.50	0.67	na	na
LA16063054	5.04	55.50	0.60	3.60	0.89	1.07	9.10	9.21	0.111	41.16	496.88	5.03	1.14	82.28	33.30	6.00	8.60	96.67	1.10	71.67	0.62	na	na
MS 2010-87-37	5.04	55.75	0.63	3.72	0.90	1.13	9.06	8.92	0.107	42.53	952.86	5.12	1.13	82.74	30.80	5.50	8.48	90.00	1.30	62.00	0.67	16.66	18.14
TAM 12J-39	5.05	45.88	0.64	4.05	0.95	1.10	11.05	10.02	0.116	41.68	870.84	5.56	1.08	82.39	33.28	5.48	8.33	90.00	1.18	78.00	0.63	20.65	20.15
TAM 13S-03	5.00	53.75	0.58	3.49	0.85	1.07	9.32	9.73	0.117	39.93	875.49	4.97	1.08	81.45	28.81	6.16	8.68	74.00	0.97	74.00	0.68	20.03	20.53
TAMLBB15905	5.04	50.13	0.63	3.61	0.88	1.12	10.08	9.67	0.095	37.56	573.04	4.80	1.19	83.70	35.46	5.70	7.88	82.86	1.45	74.29	0.84	17.52	19.56
TAMLBB16507	5.02	48.75	0.72	3.81	0.94	1.18	10.32	10.04	0.087	37.77	515.01	4.53	1.12	80.31	31.00	5.54	9.75	92.50	0.99	76.00	0.78	20.10	19.59
UA 222 CK	5.02	52.00	0.63	3.66	0.88	1.14	9.67	9.70	0.103	39.52	779.62	4.99	1.13	82.04	31.16	6.29	8.90	92.00	1.47	73.33	0.69	20.38	20.56

[z] **This data represents an average across seed grown at four different locations and was generated at the Univeristy of North Texas using the Horn et al. (2011) time domain 1H- NMR method.**

ACKNOWLEDGMENTS

We would like to thank Ted Wallace, the coordinator of the Regional Breeders Testing Network, and all the participants for providing the cotton entries; Patrick Ashe and Melissa Stefanek for taking the images and preparing the germination assays; and Kelly Thorp for help with the Python code. This work was supported by a Cotton Incorporated Research Grant, 13-738 and the United States Department of Agriculture–Agricultural Research Services, 2020-2100-013-00D.

AVAILABILITY OF MATERIALS AND DISCLAIMER

The source code for the ImageJ macro and subsequent Python code are posted in the ALARC-HTP github page, access can be granted if requested. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA. The USDA is an equal opportunity employer.

REFERENCES

Allen, R.G., L.S. Pereira, D. Raes, and M. Smith. 1998. Crop evapotranspiration–guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56. Rome: Food and Agriculture Organization of the United Nations.

Breaux, R.D. 1954. A genetic analysis of the major components of yield in American upland cotton. Ph.D. diss. Louisiana State Univ., Baton Rouge.

Bridge, R.R., W.R. Meredith, Jr., and J.F. Chism. 1971. Comparative performance of obsolete varieties and current varieties of upland cotton. Crop Sci. 11:29–32.

Campbell, B.T., P.W. Chee, E. Lubbers, D.T. Bowman, W.R. Meredith, Jr., J. Johnson, and D.E. Fraser. 2011. Genetic improvement of the Pee Dee cotton germplasm collection following seventy years of plant breeding. Crop Sci. 51:955–968. doi: 10.2135/cropsci2010.09.0545

Cook, O.F. 1908. Danger in Judging Cotton Varieties by Lint Percentage. USDA Bureau of Plant Industry Cir.11.

Cotton Incorporated. 2020. Monthly Economic Letter: Cotton Market Fundamentals and Price Outlook. Available at http://www.cottoninc.com/corporate/Market-Data/MonthlyEconomicLetter (verified 25 Aug. 2020).

Culp, T.W., and D.C. Harrell. 1975. Influence of lint percentage, boll size, and seed size on lint yield of upland cotton with high fiber strength. Crop Sci. 15(6):741–746.

Desalegn, Z., N. Ratanadilok, and R. Kaveeta. 2009. Correlation and heritability for yield and fiber quality parameters of Ethiopian cotton (*Gossypium hirsutum* L.) estimated from 15 (diallel) crosses. Kasetsart J. (Nat. Sci.) 43:1–11.

Dowd, M.K, S.M. Pelitire, and C.D. Delhom., 2018. Seed-fiber ratio, seed index, and seed tissue and compositional properties of current cotton cultivars. J. Cotton Sci. 22:60–74.

Edmisten, K. 2015. Seed size and cool germination effects on cotton stand, early growth, and yield. NC State Extension, NC State Univ. Available at https://cotton.ces.ncsu.edu/2015/04/353745/ (verified 25 August 2020).

Groves, F.E., and F.M. Bourland. 2010. Estimating seed surface area of cottonseed. J. Cotton Sci. 14:74–81.

Hinze, L.L., P.J. Horn, N. Kothari, J.K. Dever, J. Frelichowski, K.D. Chapman, and R.G Percy. 2015. Nondestructive measurements of cottonseed nutritional trait diversity in the U.S. national cotton germplasm collection. Crop Sci. 55:770–782. doi:10.2135/cropsci2014.04.0318

Hodson, E.A. 1920. Lint frequency in cotton with a method for determination. Ark. Agric. Exp. Sta. Bull. 168:3–11.

Horn, P.J., P. Neogi, X. Tombokan, S. Ghosh, B.T. Campbell, and K.D. Chapman. 2011. Simultaneous quantification of oil and protein in cottonseed by low-field time-domain nuclear magnetic resonance. J. Amer. Oil Chem. Soc. 88:1521–1529. doi:10.1007/s11746-011-1829-5

Hunsaker, D.J., E.M. Barnes, T.R. Clarke, G.J. Fitzgerald, and P.J. Pinter. 2005. Cotton irrigation scheduling using remotely sensed and FAO-56 basal crop coefficients. Trans. ASABE. 48:1395–1407. doi: 10.13031/2013.19197

Kloth, R.H., and R.B. Turley. 2010. Physiology of seed and fiber development. p. 111–122 *In* J.McD. Stewart, D.M. Oosterhuis, J.J. Heitholt, and J.R. Mauney (eds.), Physiology of Cotton. Springer, New York, NY.

Kothari, N., B.T. Campbell, J.K. Dever, and L.L. Hinze. 2015. Combining ability and performance of cotton germplasm with diverse seed oil content. Crop Sci. 56:19–29. doi:10.2135/cropsci2015.03.0166

Krieg, D.R., and S.N. Bartee. 1975. Cottonseed density: associated germination and seedling emergence properties. Agronomy J. 67:343–347.

National Cotton Council of America. 2020. World of Cotton. Available at http://www.cotton.org/econ/world/ (verified 25 Aug. 2020).

Pettigrew, W.T., and W.R. Meredith, Jr. 2009. Seed quality and planting date effects on cotton lint yield, yield components, and fiber quality. J. Cotton Sci. 13:37–47.

Smith, W.C., and G.G. Coyle. 1997. Association of fiber quality parameters and within-boll yield components in upland cotton. Crop Sci. 37:1775–1779.

Snider, J.L., G.D. Collins, J. Whitaker, K.D. Chapman, and P. Horn. 2016. The impact of seed size and chemical composition on seedling vigor, yield, and fiber quality of cotton in five production environments. Field Crops Res. 193:186–195. doi: 10.1016/j.fcr.2016.05.002