

RESEARCH

Open Access



Feasibility study of assessing cotton fiber maturity from near infrared hyperspectral imaging technique

LIU Yongliang^{1*} , TAO Feifei², YAO Haibo^{2*} and KINCAID Russell²

Abstract

Background Fiber maturity is a key cotton quality property, and its variability in a sample impacts fiber processing and dyeing performance. Currently, the maturity is determined by using established protocols in laboratories under a controlled environment. There is an increasing need to measure fiber maturity using low-cost (in general less than \$20 000) and small portable systems. In this study, a laboratory feasibility was performed to assess the ability of the shortwave infrared hyperspectral imaging (SWIR HSI) technique for determining the conditioned fiber maturity, and as a comparison, a bench-top commercial and expensive (in general greater than \$60 000) near infrared (NIR) instrument was used.

Results Although SWIR HSI and NIR represent different measurement technologies, consistent spectral characteristics were observed between the two instruments when they were used to measure the maturity of the locule fiber samples in seed cotton and of the well-defined fiber samples, respectively. Partial least squares (PLS) models were established using different spectral preprocessing parameters to predict fiber maturity. The high prediction precision was observed by a lower root mean square error of prediction (RMSEP) (< 0.046), higher R_p^2 (> 0.518), and greater percentage (97.0%) of samples within the 95% agreement range in the entire NIR region (1 000~2 500 nm) without the moisture band at 1 940 nm.

Conclusion SWIR HSI has a good potential for assessing cotton fiber maturity in a laboratory environment.

Keywords Near infrared spectroscopy, Near infrared hyperspectral imaging, Fiber maturity, Seed cotton, Partial least squares regression

Introduction

Cotton has been grown in over 100 countries/regions primarily for its natural textile fiber (Mollae et al. 2020). Cotton fiber starts from protodermal cells on the outer integument layer of cotton seeds. Fiber maturity, quantified by the degree of secondary cell wall thickening, is a key yield trait and also an essential fiber quality attribute. Maturity is linked directly to fiber breakage and entanglement (neps) during mechanical processing, yarn processing and textile performance, and dye uptake in textile products (Anthony et al. 1988; Gordon 2007; Kelly et al. 2015; Kim et al. 2019; Long et al. 2021). For example, recently Long et al. (2021) related fiber

*Correspondence:

Liu Yongliang
Yongliang.Liu@usda.gov
Yao Haibo
haibo@gri.msstate.edu

¹ USDA, ARS, Southern Regional Research Center (SRRC), Cotton Structure & Quality Research Unit, New Orleans, LA 70124, USA

² Geosystems Research Institute, Mississippi State University, Building 1021, Stennis Space Center, MS 39529, USA



This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2023. **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/>.

properties to the ring spun yarn performance on cotton from two consecutive seasons (the 2010/2011 season and the 2011/2012 season), and reported an 8.8% difference in fiber maturity and a 17.3% difference in yarn tenacity ($\text{cN}\cdot\text{tex}^{-1}$) between two seasons. Their fiber maturity on average was 0.76 from the 2010/2011 season and 0.83 from the 2011/2012 season, and this difference impacted yarn tenacity with $15.37 \text{ cN}\cdot\text{tex}^{-1}$ from the 2010/2011 season fibers and $12.92 \text{ cN}\cdot\text{tex}^{-1}$ from the 2011/2012 season fibers.

Fiber maturity is measured on cotton lint (i.e., cotton fiber after the ginning process has removed the cotton seeds) by the established procedures practiced in the cotton industry. Current-in-use fiber maturity measurements include the referenced cross-sectional image analysis microscopy (IAM) (Guo et al. 2014; Hequet et al. 2006; Thibodeaux and Evans 1986; Xu and Pourdeyhimi 1994), advanced fiber information system (AFIS) (Bradow et al. 1996; Paudel et al. 2013), and Cottonscope (ASTM 2021; Kim et al. 2020; Paudel et al. 2013; Rodgers et al. 2012, 2015). These measurements are performed under tightly controlled laboratory environmental conditions of $(21 \pm 1)^\circ\text{C}$ temperature and $(65 \pm 2)\%$ relative humidity (RH).

To perform cotton fiber maturity measurements by the above methods, the cotton fiber attached to cotton seeds (or seed cotton) have to be harvested first in the field, then they are ginned at gin facilities prior to shipping the resultant fiber samples to a fiber testing laboratory for routine analyses. Clearly, from cotton field harvesting to cotton ginning to laboratory test, fiber maturity determination is a lengthy and tedious process involving fiber shipment and conditioning that might take days to have fiber maturity available (Zumba et al. 2017), because fiber ginning (or production) facilities and laboratory quality test laboratories are not in close proximity. Hence, there is an increased interest in rapid and accurate analysis of cotton fiber maturity using low-cost (in general less than \$20 000) and small portable systems. There have been a few works in this area; most have been led by researchers at the Agricultural Research Service (ARS) of the United States Department of Agriculture (USDA) (Rodgers et al. 2010, 2017; Zumba et al. 2017). These investigators have explored the feasibility of using portable near infrared (NIR) instruments to measure fiber maturity in and outside the laboratory. The principle is based on diffused NIR spectral reflectance that occurs when NIR light interacts with fiber sample, and also on NIR spectral intensities that originate from the overtones and combinations of fundamental group vibrations (i.e., C-H, O-H, C-O) in fiber cellulose (a major component in cotton fiber). NIR method is an indirect method for fiber maturity, since it relies on IAM, AFIS, or Cottonscope maturity values for NIR applications. In these pioneer

studies, Rodgers and colleagues mainly compared portable NIR devices to bench-top instruments using well-defined and clean lint fibers, while consistently keeping a good contact between the sample and the NIR probe/detector. Depending on availability of fiber maturity reference measurements over the years, they used fiber maturity references from Fineness and Maturity Tester (FMT) measurements in 2010 (Rodgers et al. 2010) to cross-sectional IAM tests (Rodgers et al. 2017) and Cottonscope tests in 2017 (Zumba et al. 2017). In addition to lint fiber, they collected NIR spectra of individual clean and seed cotton bolls using portable NIR devices, prior to mixing the bolls for the fiber maturity reference measurement and to taking the NIR spectral average. In order to evaluate portable NIR performance, they proposed several end-state criteria that included the number of outliers $\leq 30\%$ ($\geq 70\%$ prediction method agreement between the reference method result and NIR-determined result). Their investigation showed the potential of predicting cotton fiber maturity within an agreement of ± 0.10 maturity units compared with other portable NIR measurements. It was found that the portable NIR instruments did not perform as well as bench-top NIR instruments in controlled condition and with consistent fiber contact.

Shortwave infrared hyperspectral imaging (SWIR HSI) technique has evolved as a non-destructive and rapid analytical tool for agricultural and food safety and quality inspection (Park and Lu 2015; Tao et al. 2020), because it provides not only the spectral information of any pixel in an image but also the spatial information of an image at target wavelength. In the past 10 years, SWIR HSI has found promising applications in detecting cotton trash or non-lint components and in identifying different types of raw cotton (Al Ktash et al. 2020; Ni et al. 2020; Zhang et al. 2016, 2017). To our best knowledge, SWIR HSI has not been applied to the evaluation of cotton fiber maturity, although there was a report of using multispectral sensors (or imaging) for cotton fiber micronaire property that reflects fiber maturity and fineness (Sui et al. 2008).

To address the need of applying SWIR HSI or similar sensing measurements to fiber maturity in fiber research and testing laboratories, this study compared the result of expensive NIR spectrometer on well-prepared and clean fibers to that of SWIR HSI system on locule fibers in seed cotton at a laboratory environment. Seed cotton harvested by cotton mechanical harvesters contains approximately 7% (machine-picked) and 32.5% (machine-stripped) foreign or non-lint materials (Hardin and Byler 2013), and it is a challenge to perform NIR fiber maturity analysis on seed cotton directly when foreign matter at the surface of a sample is apparent, as exemplified in Fig. 1 (a). The main objective of this study was to compare

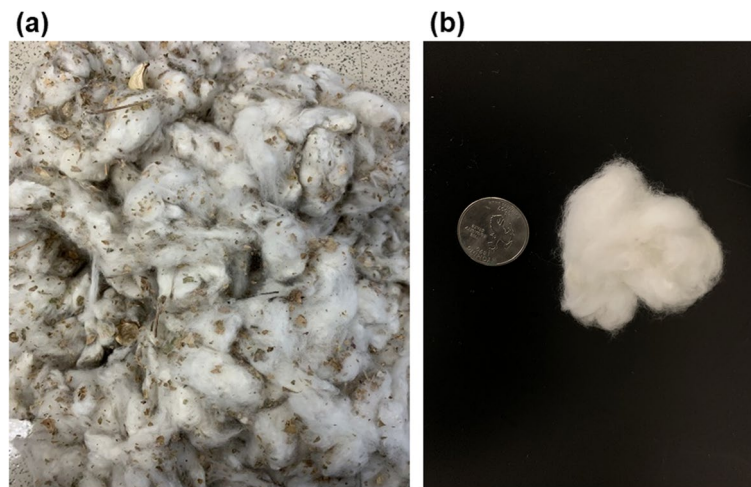


Fig. 1 **a** Example of seed cotton with the presence of non-lint materials such as broken hulls, branches, and leaf pieces that are entangled with locule fibers. **b** Selected single locule fiber sample with cotton seed attached still, but without non-lint material; a U.S. 25-cent coin (24.26 mm in diameter) was included to compare the shape and size of locule fiber sample

partial least squares (PLS) regression model performance for predicting fiber maturity from both optical NIR spectra of lint samples and SWIR HSI spectra of locule fibers still attached to their seeds at a laboratory environment. Our rationale for comparing fiber maturity analysis between NIR instrument and SWIR HSI system is:

(1) Optical NIR models were developed from well-defined fibers and their maturity values were determined following the cross-sectional IAM procedure. IAM maturity determination requires a bundle of fibers to be selected and combed, embedded in a medium, cut into thin sections, and mounted on a microscope slide, prior to acquiring and analyzing fiber cross-sectional images. Following the image analysis with a piece of dedicated software, it is possible to estimate the parameters including the fiber secondary cell wall area (A_w , cross-sectional area minus lumen area) and the perimeter of the outside of the fiber (P_2), and further to calculate fiber characteristics such as the circularity ($\theta = 4\pi A_w / (P_2)^2$), maturity ratio or maturity ($M = \theta / 0.577$) for each cross-sectional fiber as well as the average maturity and maturity distribution in a sample. In general, IAM fiber maturity test is a tedious, labor intensive, and slow process, hence it is impractical for measuring a large number of fiber samples. For the NIR spectral collection, a consistent pressure was applied to keep the fiber sample surface flat, and each NIR spectrum consisted of $\sim 2\,800$ datapoints from 1 000 to 2 500 nm.

(2) SWIR HSI models were created from locule fibers in seed cotton and their respective fiber maturity values were acquired using the Cottonscope method. Unblended cotton lint is a mixture of locule fibers at different maturity levels, due to the fiber-to-fiber maturity variation within locules in a single boll and also among

bolls from differing plant position and differing field areas (Ayele et al. 2018; Turner et al. 2015). By focusing on fiber samples from individual locules in seed cotton, it is anticipated to have an accurate locule fiber maturity value and also to exclude the impact of non-lint materials (exampled in Fig. 1 (a)) on SWIR HSI model efficiency at this stage. In the Cottonscope reference measurement, a fiber bundle was cut into snippets, then $\sim 20\,000$ individualized snippets were analyzed for the fiber maturity by using a polarized microscopy. Comparison of IAM maturity to Cottonscope maturity indicated a good overall agreement between the two methods for the 104 well-defined cottons (Paudel et al. 2013; Rodgers et al. 2012). Unlike the flat sample surface presented in the optical NIR measurement, locule fiber samples were in a fluffy state, presenting a rough surface during the SWIR HSI acquisition. A SWIR HSI spectrum includes 156 datapoints in the region of 1 000 to 2 500 nm, which is much less than that of the NIR spectrum discussed above.

Materials and methods

Well-defined cotton fibers, IAM maturity reference measurement, and NIR spectral collection

One hundred and four sub-samples from a pool of well-defined cotton fibers, known as the 104 reference materials collected by Texas Tech University for cotton fiber maturity measurements (Hequet et al. 2006) were used. They were well homogenized and clean fibers representing two main cotton types (Upland and Pima), various growth locations (U.S. domestic and foreign countries), and a wide range of fiber maturity. Maturity values (M_{IAM}) were obtained by the cross-sectional IAM protocol with the use of the FIAS software developed by XU Bugao (Hequet et al. 2006).

NIR spectra in reflectance mode were acquired on a Foss XDS rapid content analyzer (Foss NIRSystems Inc., Laurel, Maryland, USA). Approximately 10.0 g of cotton fibers were pressed into a Foss coarse granular cell with internal dimensions of 3.8 cm wide X 15.2 cm long X 4.8 cm depth. Following the granular cell moving across the optical window and stopping at 8 locations, each spectral scan covered a large surface sampling area of approximately 36 cm². To keep fiber surface flat for a good contact between the sample and the optical window during entire experiment, a 750 g of extra weight (~0.18 PSI, pound-force per square inch) was loaded on top of the fiber samples consistently. The background was recorded with the use of an internal ceramic reference tile prior to sample scan. The $\log_{10}(1/\text{Reflectance})$ or $\log_{10}(1/R)$ readings were collected over the 400~2 500 nm wavelength range at 0.5 nm interval and 32 scans at a laboratory environment. Three spectra were taken by sub-sampling different portions in a sample within 5 min and the mean spectrum was used for each sample.

Locule fibers in seed cotton

Ten commercial Upland seed cotton samples, from 8 cultivars (Deltapine 1646 (DP1646), Dyna-Gro 3385 (DG3385), Fiber Max 1830 (FM1830), Phytogen 350 (PHY350), Phytogen 430 (PHY430), Phytogen 444 (PH444), Stoneville 4848 (ST4848), and Stoneville 5471 (ST5471)), were collected at modules in the field randomly from 3 U.S. States (Mississippi, Missouri, and New Mexico) in the 2019 crop year. For each sample, 10 locules were randomly selected across a sample box containing more than 5 lbs of seed cotton, in which fibers were mingled together between locules as well as with non-lint material (Fig. 1 (a)). Following the removal of visible trash portions manually, locule fiber samples, with cotton seed presence and > 1.0 g in weight each (Fig. 1 (b)), were imaged first as detailed in the following paragraph. Before taking the images, all seed cotton locule fiber samples were conditioned for a minimum of 24 h at a laboratory environment and then scanned at this laboratory.

SWIR HSI collection, image processing, and ROI (Region of interest) spectra of locule fibers

Next, a line-scanning SWIR hyperspectral imaging system in the spectral region of 1 000~2 500 nm was used (Tao et al. 2020). It comprises an imaging spectrograph (Micro-Hyperspec® SWIR, Headwall Photonics Inc., Bolton, Massachusetts, USA), a 16-bit mercury cadmium telluride (MCT) detector coupled with a 25-mm focal length lens, halogen illumination source, and a computer. The imaging system has a spectral resolution of 9.6 nm, and the camera has a resolution of 384 (spatial dimension) X 166 (spectral dimension) pixels.

Hyperspectral images were acquired in reflectance mode from both sides of a locule fiber sample in the fluffy state. Each image took 13 s to acquire and included two locule samples with an exposure time of 10 ms and a pixel size of 778 X 384, and each pixel has a dimension of 0.273 4 mm (x-axis) and 0.254 1 mm (y-axis). Prior to the collection of fiber images, white and dark reference images were taken to correct the influence of non-uniform light intensity distribution and dark current of the imaging system. The white reference image was acquired from a white reflectance reference standard and the dark reference image was collected by covering the camera lens completely with its opaque cap. The relative reflectance (R) intensities of fiber samples were calculated from the equation: $R = (I_{\text{sample}} - I_{\text{dark}}) / (I_{\text{white}} - I_{\text{dark}})$, in which I_{sample} , I_{dark} , and I_{white} represent the intensities from sample image, dark reference image, and white reference image, respectively (Tao et al. 2020). The sample moving speed of 19 mm·s⁻¹ was calculated via the system control software.

To separate fiber pixels from background pixels in ENVI software (L3Harris Geospatial, Broomfield, Colorado, USA), a wavelength with the highest reflectance at 1316.82 nm was selected to enhance the largest contrast between the sample and the background. A band thresholding process was utilized to create a binary mask of all pixels with reflectance values of at least 0.10 at the wavelength of 1 316.82 nm. Pixels within the masked area were used to create a unique region of interest (ROI) for each of the locule sample contained within an image. After averaging two ROI reflectance spectra obtained from two sides of the individual sample, resultant mean reflectance spectra were converted to absorbance spectra for all samples before exporting them for analysis.

Locule fiber cottonscope maturity reference determination

After taking SWIR HSI scans, these locule fiber samples were measured for fiber maturity (M_{CS}) by the Cottonscope (Cottonscope® Pty Ltd, Western Australia, Australia) as described previously (Kim et al. 2020; Rodgers et al. 2012, 2015). Standard manufacturer's operational procedures were followed with the use of Cottonscope software version 2.08. All samples were conditioned for a minimum of 24 h prior to measurements and measured at standard conditions of (21 ± 1) °C temperature and (65 ± 2)% RH. Succeeding the removal of cotton seeds from each locule manually, remain fibers were cut into approximately 0.7 mm snippets using a dual-blade cutter. Then, a mass of 50.0 mg ± 0.3 mg of snippets was placed into the Cottonscope water bowl that contained 125 g of water solution with a surfactant, and snippets were individualized to begin the measurement. Each Cottonscope measurement analyzes 20 000 snippets, and a total

of 3 measurements were taken for individual locule fiber samples before utilizing the average of 3 replicates for analysis. During the experiments, the USDA Agricultural Marketing Service (AMS) micronaire calibration cotton samples (GM-39 and DM-9) were used to check instrumental performance daily (Kim et al. 2020).

NIR M_{IAM} and SWIR HSI M_{CS} model development

Both NIR spectra and SWIR HSI spectra were imported into GRAMS IQ application in Grams/AI (Version 9.1, Thermo Fisher Scientific, Waltham, Massachusetts, USA) for separate PLS regression model development. Both NIR M_{IAM} and SWIR HSI M_{CS} values were determined on 104 reference materials and 100 locule fibers, respectively, as described in previous paragraphs. On the order of the smallest to largest in M_{IAM} value, 70 spectra (or samples) were selected for calibration equation development and the remaining 34 spectra (every 3rd sample in the order) were used for model prediction. Similarly, from the smallest to largest in M_{CS} value, 67 spectra were selected for calibration equation development and 33 spectra were used for model validation. To optimize the accuracy of the prediction models, the spectra were subjected to different combinations of spectral pretreatments (e.g., mean centering (MC), multiplicative scatter correction (MSC), standard normal variate (SNV), and the first and second derivatives). Full (one-sample-out rotation) cross-prediction method was used, and

the number of optimal factors chosen for the regression equation generally corresponded to the minimum of the predicted residual error sum of squares (PRESS). The model accuracy and efficiency were evaluated in the calibration and prediction set on the basis of coefficient of determination (R_C^2 , R_P^2), root mean square error of calibration (RMSEC) and prediction (RMSEP), as well as the number (or percentage) of samples being predicted accurately. A sample was considered to be predicted accurately if the difference between NIR predicted and referenced maturity value was within the 95% agreement range of bias $\pm 1.96 \times \text{RMSEP}$ (Bland and Altman 1986). The higher R_P^2 (Chicco et al. 2021), lower RMSEP, and greater percentage of correctly predicted samples, the better and more robust the optical and imaging NIR systems response to fiber maturity. Both P -values at confidence level of 95% and F -test were calculated using respective Regression and F -Test function under Data Analysis in Microsoft® Excel® for Office 365.

Results and discussion

M_{IAM} prediction by NIR spectra

Figure 2 shows representative $\log_{10}(1/R)$ spectra (1000–2500 nm) of cotton fibers with different M_{IAM} ranges. These spectra were obtained by averaging the spectra of consecutive M_{IAM} values in the respective range of <0.70, 0.70–0.80, 0.80–0.90, 0.90–1.00, and >1.00. Spectra of cotton fibers with low M_{IAM} value have

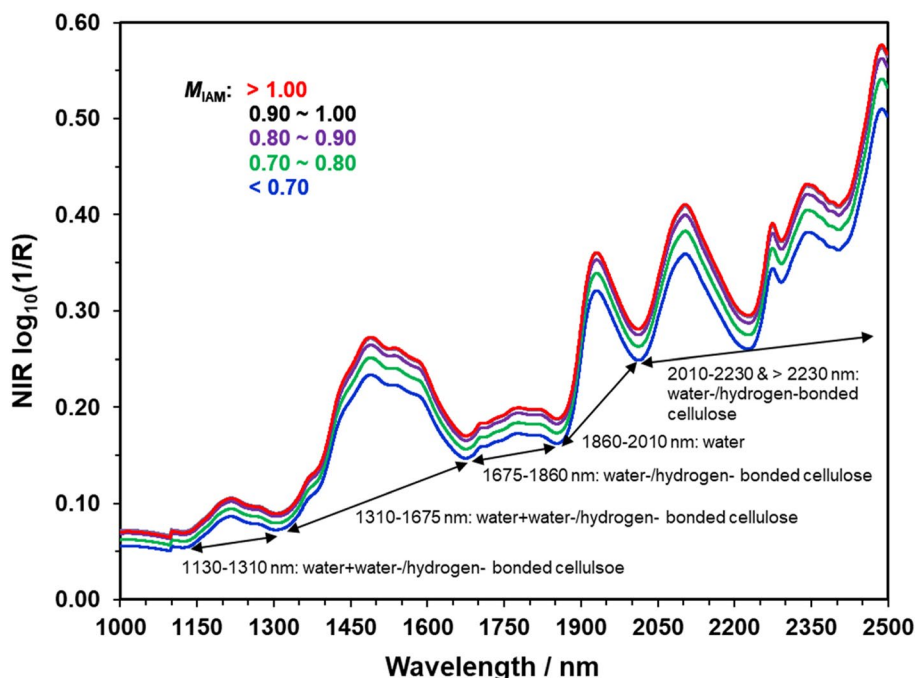


Fig. 2 NIR $\log_{10}(1/R)$ spectra of cotton fibers with various M_{IAM} values, by averaging the spectra of samples in respective M_{IAM} range of <0.70, 0.70–0.80, 0.80–0.90, 0.90–1.00, and >1.00, from bottom to top

common NIR bands with those of fibers having high M_{IAM} index, but there are overall intensity increases in entire spectral region with M_{IAM} rising. Spectral intensity variations suggest the capability of NIR technique to monitor changes in fiber maturity through quantitative multivariate data analysis. Notably, spectral difference between two mean spectra (representing $M_{IAM} = 0.90$ to 1.00 vs. $M_{IAM} > 1.00$) is indiscernible, revealing insignificant spectral intensity changes among high M_{IAM} fibers due to less variation in cellulose amount, a major fiber composition. This pattern is consistent with Fig. 4 that is from different cotton samples, different maturity measurements, and different NIR instruments. Various functional groups in fiber cellulose are responsible for characteristic bands in this region and their assignments are summarized in Table 1 (Burns and Ciurczak 2001). The weak bands in 1 150~1 300 nm region are mainly from the second overtones of CH stretching modes in water-bonded and hydrogen-bonded cellulose, while their first overtones appear in 1 675~1 860 nm region. Absorptions in 1 300~1 400 nm region are assignable to combination bands of the CH vibrations mostly, and also include the extended wing as low as 1 125 nm from the neighboring broad and the intense band at 1 450 nm. These bands in the 1 400~1 675 nm region are due to the overlap of the first overtones of the OH stretching modes

in hydrogen bonded forms from both fiber cellulose and water or moisture adsorbed in fiber. The strong bands centered at 1 940 and 2 105 nm are most likely attributed to the combination of OH stretching and deformation mode from moisture in fiber and the combination of OH and CO stretching vibrations of water-bonded and hydrogen-bonded cellulose, respectively. Intense absorptions over 2 200 nm are primarily from combination bands in water-bonded and hydrogen-bonded cellulose. Although the 1 940 nm band is the moisture band in cotton fiber, other absorption bands are a collective contribution of all chemical components in a sample, including fiber cellulose and moisture. Since moisture fluctuations may influence the NIR spectra, all samples were equilibrated in the laboratory for 24 h prior to all measurements.

Range and standard deviation (SD) of referenced M_{IAM} values in calibration and prediction sets are summarized in Table 2. Within the goal to create a set of reference cotton for maturity measurement (Hequet et al. 2006), variations of M_{IAM} values in Table 2 could represent the variability of fiber maturity in commercial cotton bales and cotton breeding programs.

PLS models were developed from different combinations of spectral pretreatments in the 1 000~2 495 nm region by cutting moisture band (1 860~2 010 nm), and the optimal result from the MC+MSC+FD spectral pretreatment is given in Table 2. This model produces an R_c^2 of 0.828 and R_p^2 of 0.739 as well as an RMSEC of 0.041 and RMSEP of 0.044. A scatter plot of NIR predicted against measured M_{IAM} in prediction set is given in Fig. 3A ($P < 0.001$). *F*-test results showed a non statistically significant difference in the variances of predicted and measured M_{IAM} index ($P = 0.19$). Meanwhile, Bland–Altman plot in Fig. 3B showed the visualization of individual samples in comparison with the 95% limits of agreement between the predicted and referenced (or actual) M_{IAM} property in the prediction set. The 95% agreement ranges are bias $\pm 1.96 \times \text{RMSEP}$, in which the

Table 1 NIR band assignment of cotton fibers (s = strong, w = weak)

Wavelength / nm	Assignment
1 150~1 300 (w)	2 nd overtone of CH stretching modes
1 300~1 400 (w)	Combination bands of CH vibrations
1 400~1 675 (s)	1 st overtone of OH stretching modes
1 675~1 860 (w)	1 st overtone of CH stretching modes
1 940 (s)	Combination of OH stretching and deformation modes
2 105 (s)	Combination of OH and CO stretching vibrations
2 270 (s), 2 340 (s)	Combination bands

Table 2 Statistics in calibration and prediction sets for M_{IAM} or M_{CS} property from NIR or SWIR HSI spectra

Maturity & optimal factor ^a	Calibration set					Prediction set				
	Range	SD ^b	R_c^2	RMSEC ^b	No. / % ^c	Range	SD ^b	R_p^2	RMSEP ^b	No. / % ^c
M_{IAM} , 5	0.570~1.087	0.099	0.828	0.041	68 / 97.1	0.632~1.030	0.085	0.739	0.044	33 / 97.1
M_{CS} , 8	0.530~0.970	0.076	0.835	0.036	66 / 98.5	0.650~0.947	0.064	0.518	0.046	32 / 97.0

^a Optimal models from spectral preprocessing with combinations of mean centering (MC), multiplicative scatter correction (MSC), and Savitzky-Golay (SG) first derivative (FD) function of thirteen points; For M_{IAM} , 70 samples in calibration set and 34 samples in prediction set; For M_{CS} , 67 samples in calibration set and 33 samples in prediction set.

^b Standard deviation (SD), root mean square error of calibration (RMSEC) and prediction (RMSEP).

^c Number and percentage (%) of samples with the difference between predicted and referenced value within the 95% agreement range of bias $\pm 1.96 \times \text{RMSEC}$ (or RMSEP) in calibration and prediction set.

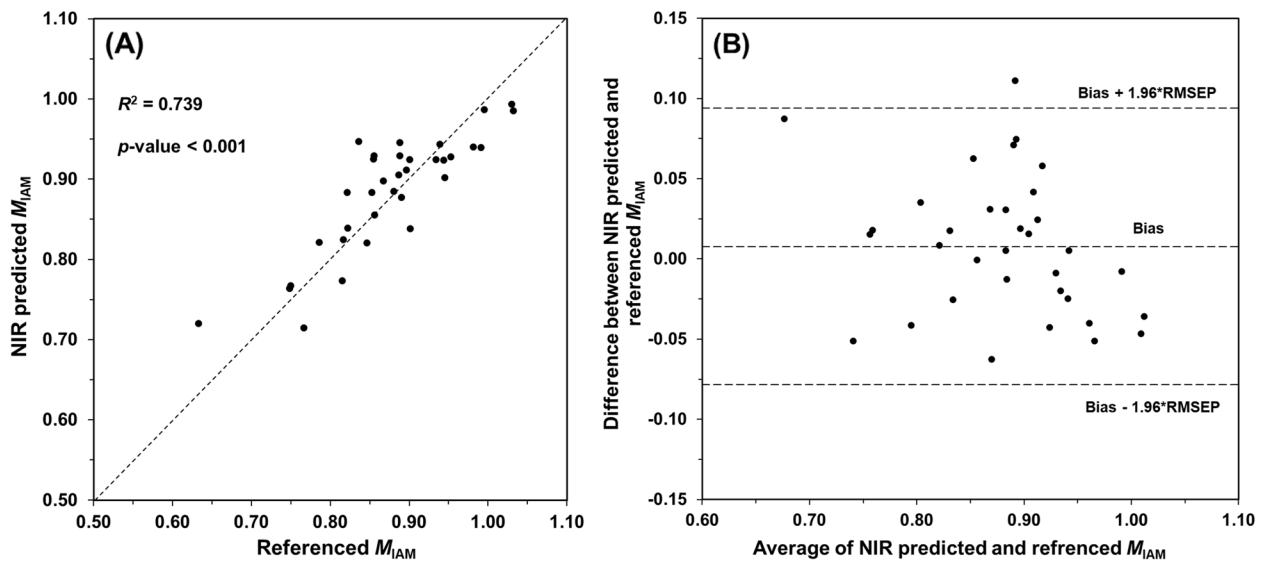


Fig. 3 A Comparative correlation between NIR predicted and referenced M_{IAM} in prediction set ($n=34$) from 1 000~2 495 nm without the moisture 1 940 nm band; (B) Bland-Altman plot of NIR predicted and referenced M_{IAM} property

bias is the difference between the averages of the predicted and referenced M_{IAM} values while RMSEP is the root mean square error of the prediction. The number of outliers, i.e., the samples outside the agreement ranges, is 1, indicating that around 97.0% ($((34-1)/34)*100\%$) of validation samples are in agreements at $\alpha=0.05$. Similarly, the number of outliers in calibration samples is 3, meaning an over 97.0% agreement.

M_{CS} prediction by SWIR HSI spectra

As expected, SWIR HSI $\log_{10}(1/R)$ spectra (1 000~2 500 nm) of locule fibers with cotton seed presence in Fig. 4 resemble NIR $\log_{10}(1/R)$ spectra of cotton fibers in Fig. 2 well. Figure 4 shows characteristic cotton fiber NIR bands from 1 350 to 1 650 nm, 1 875 to 2 000 nm, 2 020 to 2 200 nm, and 2 240 to 2 400 nm with absorption peaks at 1 485/1 560, 1 940, 2 105, and 2 275/2 345 nm.

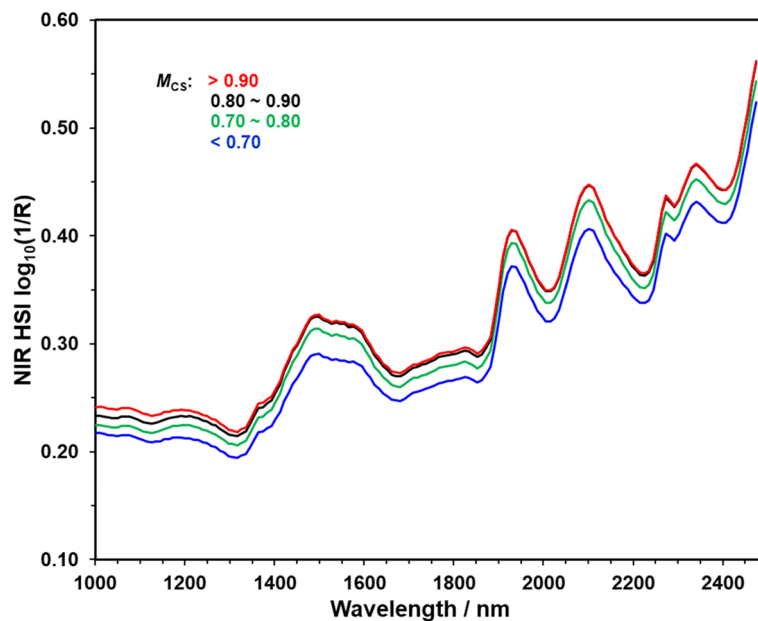


Fig. 4 SWIR HSI $\log_{10}(1/R)$ spectra of locule fibers with cotton seed presence in different M_{CS} intervals, by averaging the spectra of neighboring M_{CS} values in respective range of < 0.70 , $0.70 \sim 0.80$, $0.80 \sim 0.90$, and > 0.90 , from bottom to top

However, the difference in baselines or baseline shift of two types of NIR spectra is distinctive between Fig. 2 (<0.10 at 1 100 nm) and Fig. 4 (>0.20 at 1 100 nm). This shift was not unexpected, because optical NIR spectra were acquired from a flat sample surface, whereas SWIR HSI spectra were collected from a fluffy fiber sample with a rough surface. Nevertheless, these NIR spectra reflect the diffused reflectance from fiber samples' surface. Effect of the baseline shifts on spectral modeling power can be minimized significantly by trying different several statistical analysis strategies, such as derivatives and MSC. Similarity of the NIR curve in Fig. 4 to those in Fig. 2 suggests that the SWIR HSI technique has an equivalent power detecting cotton fiber maturity as the NIR.

Compared with the range (0.570 to 1.087) and SD (0.099) within calibration samples for the 104 reference materials in Table 2, both the range (0.530 to 0.970) and SD (0.076) of calibration set for the 100 locule fiber samples in Table 2 decreased. For example, the low-end of range changes to 0.530 for the 100 locule fibers from 0.570 for the 104 reference materials, the high-end of range becomes 0.970 for the 100 locule fibers from 1.087 for the 104 reference materials, and SD reduces to 0.076 for the 100 locule fibers from 0.099 for the 104 reference materials. This is not surprising, since the 104 reference samples were created subjectively to cover a large maturity range (Hequet et al. 2006). For the 104 reference materials, Paudel et al. (2013) showed a dynamic M_{CS} range of 0.47 to 1.11 as compared with the

M_{IAM} range of 0.57~1.09, indicating a difference of 0.10 ($M_{IAM} - M_{CS}$) between the lowest M_{IAM} and the lowest M_{CS} and also of -0.02 ($M_{IAM} - M_{CS}$) between the highest M_{IAM} and the highest M_{CS} . M_{CS} on the 100 locule fiber samples were observed to be lower than M_{IAM} on the 104 reference materials in Table 2, echoing the M_{CS} and M_{IAM} observation on the same 104 reference materials by Paudel et al. (2013). Even on partial (93) of this 104 reference materials, Rodgers et al. (2012) reported a moderate R^2 of 0.65 and a slight bias (standard deviation of differences (SDD)=0.062) between M_{CS} and M_{IAM} , however, they did not specify the M_{CS} or M_{IAM} range. The authors (Rodgers et al. 2012) further ascribed these discrepancies to (i) sample inhomogeneities in natural cotton fiber and (ii) errors in M_{IAM} determination (limited number of specimens, image detection bias, fiber bundle sectioning). These researchers also noted the lower M_{CS} values for samples at higher M_{IAM} values, due to errors in M_{IAM} determination of immature versus dead fibers. Finally, they concluded a very good method agreement between Cottonscope and IAM method for fiber maturity. Despite of different fiber samples, the pattern of fiber M_{IAM} frequency (%) on the 104 reference materials is similar to that of fiber M_{CS} frequency (%) on 100 locule fiber samples (Fig. 5), implying that random selection of the 100 locule fiber samples from 10 cotton cultivars in this study is appropriate.

To correlate SWIR HSI spectra with corresponding M_{CS} values, different combinations of ROI spectral pretreatments in the 1 001~2 474 nm region without

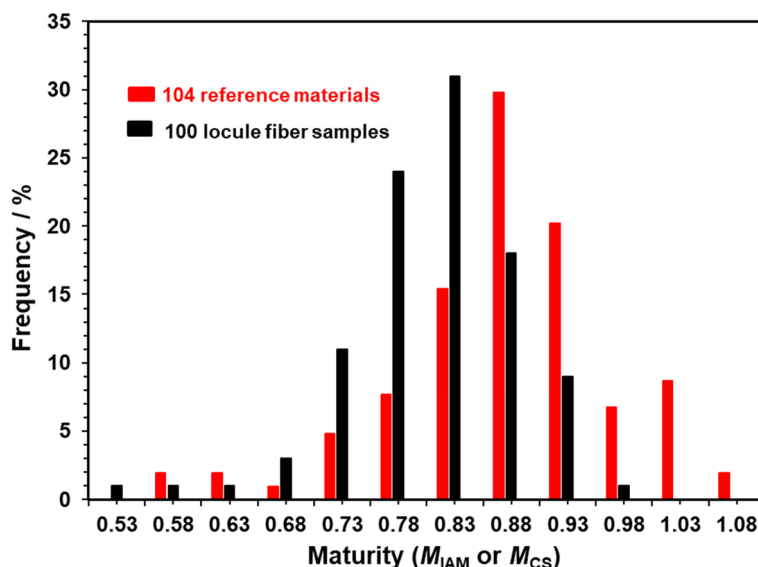


Fig. 5 Comparison of fiber maturity frequency (%) on the sample set of 104 reference materials (■) against 100 locule fiber samples (■). Curve was obtained for each sample set, after classifying the M_{IAM} or M_{CS} values into 20 bins (bin width, 0.05) and counting the frequency number of each bin

the moisture 1 940 nm band (1 860~2 010 nm) were attempted and the optimal result is inserted in Table 2.

for comparison. This model suggests an overall agreement of $\geq 97.0\%$ in 67 calibration and 33 prediction samples that are within the 95% agreement ranges, and also an RMSEC of 0.031 and RMSEP of 0.046. Figure 6 A and 6B show a correlation plot between SWIR HSI predicted against actual M_{CS} ($P < 0.001$) and a Bland–Altman plot in prediction set, respectively. Difference in the variances of predicted and measured M_{CS} value was not statistically significant ($P = 0.30$).

It is of interest to examine the M_{CS} variation at a single locule level within a cultivar (or a sample). Table 3 summarizes the M_{CS} average, standard deviation (SD), and range on 10 samples, showing a variation of M_{CS} average from 0.730 to 0.891 for 10 samples and a M_{CS} range from 0.530 to 0.970 at single locule level among 10 samples. As expected in Table 3, samples with similar M_{CS} average could show different maturity ranges. For example, the pair of PHY350 vs. PHY444 revealed a similar M_{CS} average of 0.831 vs. 0.828 but a different M_{CS} range of 0.747 to 0.893 and 0.684 to 0.970, respectively. M_{CS} frequency in Fig. 7 suggested a clear difference between the two cultivars, with a maximum M_{CS} frequency at $M_{CS} = 0.82$ and a narrow M_{CS} interval for PHY350 cultivar and for PHY444, at $M_{CS} = 0.73$ and a wide M_{CS} interval. Comparatively, DP1646 dryland and PHY430 showed a common M_{CS} mean (0.785 vs. 0.794) and M_{CS} range (0.733 to 0.827 vs. 0.733 to 0.833), and also a similar M_{CS} frequency.

Table 3 Comparison of M_{CS} avg $\pm$ SD and range on 10 seed cotton samples from cottonscope measurement

Location / cultivar		Avg $\pm$ SD	Range
Mississippi	DP1646	0.850 $\pm$ 0.021	0.827~0.893
	ST4848	0.745 $\pm$ 0.083	0.530~0.817
	ST5471	0.862 $\pm$ 0.034	0.810~0.915
Missouri	DP1646	0.802 $\pm$ 0.027	0.753~0.843
	DP1646 dryland	0.785 $\pm$ 0.031	0.733~0.827
	PHY350	0.831 $\pm$ 0.047	0.747~0.893
New Mexico	PHY430	0.794 $\pm$ 0.031	0.733~0.833
	DG3385	0.730 $\pm$ 0.082	0.570~0.833
	FM1830	0.891 $\pm$ 0.043	0.800~0.947
	PHY444	0.828 $\pm$ 0.098	0.684~0.970

Data in Table 2 indicated consistent prediction abilities in the aspect of RMSEP (0.044 from NIR optimal model and 0.046 from SWIR HSI model) and percentage of samples within the 95% agreement range ($> 97.0\%$ from NIR and SWIR HSI model) between optical NIR and SWIR HSI system, despite the facts that (1) different types of reference values were taken (M_{IAM} vs. M_{CS}), (2) different spectral datapoints or instrumental resolution were used (2 750 datapoints in optical NIR vs. 155 in SWIR HSI), and (3) different sampling methods were adopted (flat surface in optical NIR vs. fluffy fibers in SWIR HSI). On the other hand, a lower R_p^2 value (0.518) from SWIR HSI model than that (0.739) from NIR model might be due to a reduced SD value

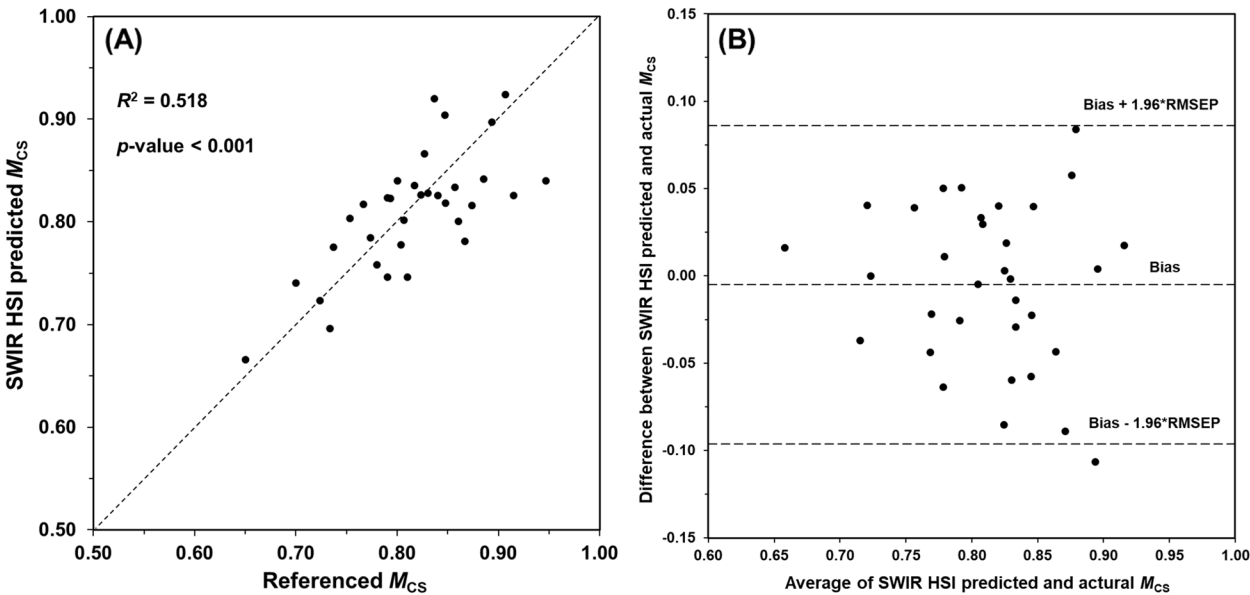


Fig. 6 **A** Comparative correlation between SWIR HSI predicted and measured M_{CS} in prediction set ($n = 33$) from 1 001~2 474 nm without the moisture 1 940 nm band. **B** Bland–Altman plot of SWIR HSI predicted and referenced M_{CS} property

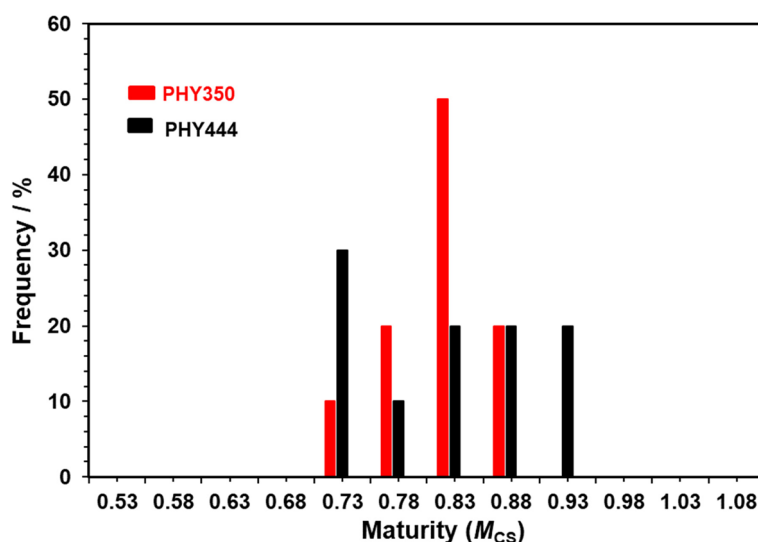


Fig. 7 Comparison of measured M_{CS} frequency (%) between PHY350 (■) and PHY444 (■)

(0.085 in NIR models vs. 0.064 in SWIR HSI models). Like the 104 reference materials developed for fiber maturity measurement, more diverse cotton cultivars from different locations and crop years are necessary to validate the SWIR HSI method. Therefore, result reveals a potential of SWIR HSI instrument for rapid and accurate measurements of fiber maturity at a laboratory environment, in which cotton samples were conditioned to the same equilibrium of both temperature and RH prior to be imaged under this condition. Also, the surface status of seed cotton should be considered in future work.

Both fiber M_{IAM} and M_{CS} models were established on conditioned samples using the 1 000~2 495 nm NIR region excluding the 1 860~2 010 nm moisture absorptions, clearly this NIR region is the minimum of the moisture effect but impacted by water-bonded cellulose. An earlier investigation (Vogt et al. 2011) indicated that exclusion of the 1 935 nm moisture band (or cutting out the 1 900~2 000 nm region) had almost no impact on fiber micronaire (an indicator of fiber maturity and fineness component) prediction. In a separate study of 30 commercial cotton bales (Knowlton 1996), result revealed the equilibrium moisture content (EMC) ranging from just under 6.7% to just over 7.2% by the oven-dry method, and also showed relatively low EMCs for high micronaire cotton at a controlled condition of 70 °F (or 21 °C) and 65% RH. Relative to a small change in EMC value from 6.7% to 7.2% (a change of 7.5%), micronaire varied largely from 3.50 to 5.20 unit (a change of 48.6%). This reported data suggested a less change in moisture content within samples at controlled environment.

Conclusion

Result of SWIR HSI system on fiber Cottonscope maturity reference of locule fiber samples in seed cotton was compared with that of a bench-top NIR instrument on fiber IAM maturity reference of well-defined cotton fiber samples. NIR spectral consistence was observed between two different spectral measurements, with common NIR bands and shapes apparent. Despite the differences in maturity reference measurements, spectral datapoints, and sampling methods between optical NIR and SWIR HSI techniques, satisfactory, consistent and comparable prediction results were observed on conditioned samples, with RMSEP < 0.044 from NIR optimal model and < 0.046 from SWIR HSI models as well as the percentage of correctly predicted samples > 97.0% at $\alpha=0.05$ from NIR and SWIR HSI models. These results indicated a potential of low-cost SWIR HSI technique in estimating fiber maturity rapidly and non-destructively at a laboratory. Further research is needed to apply the method outside the laboratory and to transfer the models between sample types as well as between instruments.

Acknowledgements

Authors wish to acknowledge Dr Thibodeaux D in providing the 104 reference materials information, Dr Delhon C in providing seed cotton samples, and Ms King H and Moraitis J in supporting Cottonscope measurement. Authors also thank Prof Xu B, Prof Li C, Dr Delhom C, Dr Pelletier M, and Dr Funk P for professional insight to improve the manuscript. Mention of a product or specific equipment does not constitute a guarantee or warranty by the U.S. Department of Agriculture and does not imply its approval to the exclusion of other products that may also be suitable.

Authors' contributions

Liu Y and Yao H conceived the research. Liu Y was responsible for all cotton samples, optical NIR and Cottonscope measurements, and Kincaid R performed SWIR data collection. Liu Y, Tao F, and Yao Y conducted data

analysis and prepared the manuscript. All authors read and approved the final manuscript.

Funding

This research was supported partially by the USDA-ARS Research Project # 6054-44000-080-00D.

Availability of data and materials

All data used or generated are included in the manuscript.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

Authors do not have competing interests.

Received: 16 August 2023 Accepted: 27 October 2023

Published online: 27 November 2023

References

- Al Ktash M, Hauler O, Ostertaga E, et al. Ultraviolet-visible/near infrared spectroscopy and hyperspectral imaging to study the different types of raw cotton. *J Spectr Imaging*. 2020;9:a18. <https://doi.org/10.1255/jsi.2020.a18>.
- Anthony WS, Meredith WR, Williford JR. Neps in ginned lint: the effects of cultivars, harvesting, and ginning practices. *Text Res J*. 1988;58:633–40.
- ASTM D8394-21. Standard test method for automated measurement of maturity, fineness, ribbon width, and micronaire of cotton fibers. West Conshohocken, PA, USA: ASTM International; 2021.
- Ayele AG, Kelly BR, Hequet EF. Evaluating within-plant variability of cotton fiber length and maturity. *Agron J*. 2018;110:47–55. <https://doi.org/10.2134/agronj2017.06.0359>.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;327:307–10.
- Bradlow JM, Hinojosa O, Wartelle LH, et al. Applications of AFIS fineness and maturity module and x-ray fluorescence spectroscopy in fiber maturity evaluation. *Text Res J*. 1996;66:545–54.
- Burns A, Ciurczak EW, editors. *Handbook of Near-Infrared analysis*. New York, USA: Marcel Dekker; 2001.
- Chicco D, Warrens MJ, Jurman G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Comput Sci*. 2021;7:e623. <https://doi.org/10.7717/peerj-cs.623>.
- Gordon S. Cotton fiber quality. In: Gordon S, Abidi N, editors. *Cotton: Science and technology*. Cambridge, UK: Woodhead Publishing Limited; 2007. p. 68–100.
- Guo X, Ouyang W, Xu B. Assessing cotton maturity using distributional parameters of fiber cross-section measurements. *Text Res J*. 2014;84:1666–76. <https://doi.org/10.1177/0040517514527375>.
- Hardin RG, Byler RK. Evaluation of seed cotton cleaning equipment performance at various processing rates. *Appl Eng Agric*. 2013;29:637–47. <https://doi.org/10.13031/aea.29.9920>.
- Hequet E, Wyatt B, Abidi N, et al. Creation of a set of reference material for cotton fiber maturity measurements. *Text Res J*. 2006;76:576–86.
- Kelly B, Abidi N, Ethridge D, et al. Fiber to fabric. In: Fang DD, Percy RG, editors, et al., *Cotton*. 2nd ed. Madison, WI, USA: American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America; 2015. p. 665–744.
- Kim HJ, Delhom CD, Rodgers JE, et al. Effect of fiber maturity on bundle and single-fiber strength of upland cotton. *Crop Sci*. 2019;59:115–24. <https://doi.org/10.2135/cropsci2018.05.0324>.
- Kim HJ, Delhom CD, Fang DD, et al. Application of the cottonscope for determining fiber maturity and fineness of an upland cotton MAGIC population. *Crop Sci*. 2020;60:2266–79. <https://doi.org/10.1002/csc.20197>.
- Knowlton JL. Effect of moisture on cotton fiber strength. Proceedings of the Beltwide Cotton Conference-2. Memphis, TN, USA: National Cotton Council; 1996. p. 1300–5.
- Long RL, Delhom CD, Bange MP. Effects of cotton genotype, defoliation timing and season on fiber cross-sectional properties and yarn performance. *Text Res J*. 2021;91:1943–56. <https://doi.org/10.1177/0040517521992769>.
- Mollaee M, Mobli A, Mutti NK, et al. Challenges and opportunities in cotton production. In: Jabran K, Chauhan BS, editors. *Cotton production*. Hoboken, New Jersey, USA: John Wiley & Sons Ltd; 2020. p. 371.
- Ni C, Li Z, Zhang X, et al. Online sorting of the film on cotton based on deep learning and hyperspectral imaging. *IEEE Access*. 2020;8:93028–38. <https://doi.org/10.1109/ACCESS.2020.2994913>.
- Park B, Lu R, editors. *Hyperspectral imaging technology in food and agriculture*. New York, USA: Springer; 2015.
- Paudel D, Hequet E, Abidi N. Evaluation of cotton fiber maturity measurements. *Ind Crop Prod*. 2013;45:435–41. <https://doi.org/10.1016/j.indcrop.2012.12.041>.
- Rodgers J, Montalvo JG, Gayle D, et al. Near infrared measurement of cotton fiber micronaire, maturity and fineness – a comparative investigation. *Text Res J*. 2010;80:780–93. <https://doi.org/10.1177/0040517509343780>.
- Rodgers J, Delhom C, Fortier C, et al. Rapid measurement of cotton fiber maturity and fineness by image analysis microscopy using the Cottonscope. *Text Res J*. 2012;82:259–71. <https://doi.org/10.1177/0040517511431317>.
- Rodgers J, Naylor GRS, Cui X, et al. Cottonscope fiber maturity, fineness, and ribbon width measurements with different sample sizes. *Text Res J*. 2015;85:897–911. <https://doi.org/10.1177/0040517514548809>.
- Rodgers J, Zumb J, Fortier C. Measurement comparison of cotton fiber micronaire and its components by portable near infrared spectroscopy instruments. *Text Res J*. 2017;87:57–69. <https://doi.org/10.1177/0040517515622153>.
- Sui R, Thomasson JA, Ge Y, et al. Multispectral sensor for in-situ cotton fiber quality measurement. *Trans ASABE*. 2008;51:2201–8.
- Tao F, Yao H, Hruska Z, et al. A novel hyperspectral-based approach for identification of maize kernels infected with diverse *Aspergillus flavus* fungi. *Biosyst Eng*. 2020;200:415–30. <https://doi.org/10.1016/j.biosystemseng.2020.10.017>.
- Thibodeaux DP, Evans JP. Cotton fiber maturity by image analysis. *Text Res J*. 1986;56:130–9.
- Turner C, Sari-Sarraf H, Hequet E, et al. Variation in maturity observed along individual cotton fibers using confocal microscopy and image analysis. *Text Res J*. 2015;85:867–83. <https://doi.org/10.1177/0040517514551466>.
- Vogt F, Luttrell R, Rodgers J. New approaches for field analyses of cotton quality by means of near-IR spectroscopy supported by chemometric. *Anal Lett*. 2011;44:2466–77. <https://doi.org/10.1080/00032719.2011.551855>.
- Xu B, Pourdeyimi B. Evaluating maturity of cotton fibers using image analysis: definition and algorithm. *Text Res J*. 1994;64:330–5.
- Zhang R, Li C, Zhang M, et al. Shortwave infrared hyperspectral reflectance imaging for cotton foreign matter classification. *Comput Electron Agric*. 2016;127:260–70. <https://doi.org/10.1016/j.compag.2016.06.023>.
- Zhang M, Li C, Yang F. Classification of foreign matter embedded inside cotton lint using shortwave infrared (SWIR) hyperspectral transmittance imaging. *Comput Electron Agric*. 2017;139:75–90. <https://doi.org/10.1016/j.compag.2017.05.005>.
- Zumba J, Rodgers J, Indest M. Fiber micronaire, fineness, and maturity predictions using NIR spectroscopy instruments on seed cotton and cotton fiber, in and outside the laboratory. *J Cotton Sci*. 2017;21:247–58.