The *Journal of Cotton Science 16:272–281 (2012)*
http://journal.cotton.org, © The Cotton Foundation 2012

TEXTILE TECHNOLOGY

Relationship Between Three Cotton Trash Measurements and Near Infrared Spectral Response

Yongliang Liu*, Devron Thibodeaux, and Gary R. Gamble

ABSTRACT

Presence of non-lint materials (trash) in commercial cotton bales at various amounts degrades market values and further influences end-use qualities. To ensure fair trading, the USDA Agricultural Marketing Service introduced the High Volume Instrument (HVI) measurement as a universal standard index. Trash contents from HVI module represent the trash portion only on a sample's surface. In additional to HVI's geometric method, gravimetric-based Shirley Analyzer (SA) and Advanced Fiber Information System (AFIS) also have been utilized to quantify the trash contents. With the increasing acceptance of HVI readings in the domestic and international trading, there is interest in the relationship between HVI and SA trash from cotton customers and regulators. Due to the complexity of not only trash type, size, and weight distribution but also the nature of HVI and SA tests, there have been few studies that compare the two types of trash readings. This study investigated the correlations between two HVI trash readings, and revealed a general conversion of HVI_{count}=104.5*HVI_{area} among low trash samples. Then, correlations between the HVI and SA trash and also against AFIS trash were examined, and a stronger relationship between HVI and SA trash than between HVI and AFIS trash was observed. Next, the samples were subgrouped subjectively according to the ratios of $HVI_{area}/SA_{visible}$ (or $HVI_{count}/SA_{visible}$), and from the plots with the lowest axis intercept values, it was proposed two general conversions of $SA_{visible}$ = 6.82*HVI_{area} and $SA_{visible}$ = 0.069*HVI_{count}. To verify the likely conversion, NIR spectra were correlated with HVI_{area} readings.

Y. Liu*, USDA, ARS, Cotton Structure & Quality Research Unit, USDA-ARS, Southern Regional Research Center, P.O. Box 19687, New Orleans, LA 70124; D. Thibodeaux, 103 Long View Ct, Pickens, SC; and G. Gamble, USDA, ARS, Quality and Safety Assessment Research Unit, 950 College Station Rd, Athens, GA
*Corresponding author: yonglian.liu@ars.usda.gov

Because of economic factors, virtually the entire cotton crop in the U. S. and Australia is harvested by machine (Wakelyn et al., 2007). Mechanically harvested cotton contains 13 to 35% foreign matter or plant-related contaminants and other irregularities (Funk et al., 2005). Considerable efforts have been made to remove the foreign matter (or trash) during the ginning and cleaning practices (Anthony, 2007). However, it is impossible to separate all trash from lint fibers even with the implementation of cleaners at the ginning sites.

Presence of trash in various amounts degrades the market value and also influences the end use for yarn and fabric processing. Both human classer inspection and High Volume Instrument (HVI) methods have been regulated by the USDA Agricultural Marketing Service (AMS) for leaf grade classification and determination of trash content in lint cottons (USDA, 2005). To ensure fair trading, USDA AMS introduced HVI measurements as indices of universal fiber quality for global implementation (Knowlton, 2002). HVI is an automatic testing procedure that within 1 to 2 min. can provide several important fiber characteristics simultaneously, such as micronaire, strength, length, short fiber index, uniformity index, and color. HVI trash contents are generated by one of three HVI modules and represent the trash portion only on the sample's surface.

Extensive studies have been undertaken to develop a number of physical and optical instruments for measuring trash content (Gordon, 2007). The instruments can be categorized into two main methods: gravimetric based and geometric (or surface scanner) based. At present, representatives in the gravimetric group are Shirley Analyzer (SA) and Advanced Fiber Information System (AFIS), which separate the trash components by mechanical means and then collect information by weighing; the geometric group includes the current HVI lines and imaging devices (Xu et al., 1997), which perform optical surface scanning. Each of these testing methods has unique advantages and limitations, and the comparison among them is summarized in Table 1. For example, SA and AFIS are destructive during the analytical process, and the tested samples cannot be reused. In general, the SA procedure is labor-intensive and time-consuming (~15 min. for one sample), whereas AFIS is rapid (~2 min.)

but requires careful sample preparation. On the other hand, HVI can assess only the trash content in terms of particle count and percentage area of the sample's surface and, therefore, does not yield any information about the weight of trash within bulky samples. In addition, AFIS and HVI provide fiber properties other than trash contents, and SA has a small-scale cleaning function that makes it a useful tool for determining visible and invisible trash content in cottons.

Table 1. Comparison of three cotton trash measurement methods.

	HVI	SA	AFIS
type	non-destructive	destructive	destructive
configuration	geometric	gravimetric	gravimetric
amount (g)	10	100	0.5
time (min.)	2	15	2
trash reading	particle & area	visible & invisible	visible foreign matter
other readings	yes		yes
cleaning effect		yes	
preparation	yes		yes

Because of the heterogeneous distribution of trash type and size, the use of different sampling specimens during three independent measurements, and the availability of all three instruments (especially SA) at any one cotton fiber research facility, there are few literature studies that compare trash readings among the three methods. In addition, because of relatively small sample size (~0.5 g) used in AFIS procedure, trash readings from HVI and SA have been frequently cited. Because USDA AMS has regulated the HVI trash index as a global fiber quality characteristics, there is interest from domestic and foreign customers to relate the geometric-based HVI trash readings with the gravimetric-based SA values. Given the complexity of trash in lint fibers and the nature of HVI and SA measurements, it is a challenge to unravel the relationship between the two types of trash determination.

Near infrared (NIR) spectroscopy, a useful technique due to the speed, ease of application, and adaptability to on-line or off-line implementation, has been used for the prediction of trash contents either from HVI reading (Liu et al., 2010a; Thomasson and Shearer, 1995) or SA and AFIS procedures (Liu et al., 2010a, 2010b). For example, Thomasson and Shearer (1995) reported the optimal NIR models for eight cotton quality characteristics and observed the lowest R^2 value (0.60) for HVI trash components. In a recent study evaluating three trash measurements by NIR technique, trash models built from HVI particle, HVI area, AFIS, and SA references in the 1100 to 2500 nm region were reported to have an R^2 of 0.80, 0.69, 0.82, and 0.82, respectively (Liu et al., 2010a). Even though the UV/visible/NIR models on visible trash and cotton fiber content in cotton waste were slightly improved (R^2 = 0.86) with SA readings in the 1100 to 2496 nm region, these models still show the difficulty of precise and quantitative determination of visible trash and cotton fiber portions for quality control purposes (Liu et al., 2010b). Hence, a 90% confidence interval was implemented to remove outlier samples that exhibited larger differences between NIR predicted and measured references, and the recalibrated models revealed the feasibility of NIR technique in the determination of trash content (Liu et al., 2010a, 2010b). Our strategy to exclude the outliers in developing reliable and robust NIR models is reasonable because of (1) high diversification of trash types and their heterogeneous distribution, (2) lint fiber mingled with visible trash and likely resulting in errors during SA reference determination (Montalvo and Mangialardi, 1983), (3) near-surface characterization (~2.5 mm) of bulky samples in spectral reflectance acquisition (Haanstra et al., 1998), and (4) varying sampling specimens between NIR spectral and reference measurements.

The objectives of this study were (1) to examine the correlation between two HVI trash readings, particle count (HVI_{count}) against percentage area (HVI_{area}); (2) to correlate the HVI trash with visible trash content ($SA_{visible}$) from SA; (3) to relate the HVI trash with visible foreign matter content ($AFIS_{VFM}$) from AFIS; (4) to subgroup the samples with the $HVI_{area}/SA_{visible}$ ratios and further explore the relationship between HVI_{area} and $SA_{visible}$ readings for each subset; and (5) to verify the proposed relationship from independent NIR spectral response with the aid of partial least square (PLS) modeling.

MATERIALS AND METHODS

Cotton Samples. A total of 406 lint cottons were collected over a 4-yr span at the ARS Cotton Quality Research Station (Clemson, SC) to represent diverse distributions in cotton variety, growing years, and locations. These samples were from normal cotton bales and contained low to medium levels of trash. They were well conditioned at a constant relative humidity of 65% and temperature of 22 ± 2 ºC prior to trash content measurements and visible/NIR spectral collection.

Determination of Cotton Trash Contents. Cotton trash contents were measured in four reference indices by three methods, namely, visible trash content ($SA_{visible}$, %) from SA (Shirley Developments,

Ltd., Stockport, UK), visible foreign matter content ($AFIS_{VFM}$, %) from AFIS (Uster Technologies, Inc., Charlotte, NC), and HVI particle count (HVI_{count}) and percentage area (HVI_{area}, %) from HVI (Uster Technologies, Inc., Charlotte, NC). Following standard procedures (ASTM, 2007a, 2007b, 2007c), approximately 100, 0.5, and 10 g of lint cottons from different fractions of the same cotton bale were processed for trash content by SA, AFIS, and HVI, respectively. Two readings from SA and HVI measurements and three replicates from AFIS measurement were averaged and analyzed further. To keep the common usage of trash readings from HVI, SA, and AFIS, direct numbers (instead of %) were used in this study. For example, the 3.65 value represented a sample with HVI area (or SA or AFIS) trash of 3.65%.

Visible/NIR Reflectance Measurement. Visible/NIR reflectance spectra were acquired on a Foss XDS rapid content analyzer (Foss NIRSystems Inc., Laurel, MD). Approximately 10 g of cotton fibers were pressed into a Foss coarse granular cell, which is rectangular with internal dimensions of 3.8-cm wide x 15.2-cm long x 4.8-cm deep. To keep good contact between the cotton sample and optical window, 750 g of extra weight was loaded on the top of fiber samples consistently throughout the entire experiment. Background was recorded with the use of an internal ceramic reference tile before scanning the samples. The log (1/Reflectance) readings were acquired over the 400- to 2500 nm wavelength range at 0.5 nm interval and 32 scans. Three spectra were collected for each of the cotton samples by repacking and the mean spectrum was used for model development.

Partial Least Squares (PLS) Models. The visible/NIR spectra were imported into the PLSplus/IQ package in Grams/AI (Version 7.01, Galactic Industrious Corp., Salem, NH [now part of Thermo Fisher Scientific]) and were smoothed with a Savitzky-Golay function (polynomial = 2 and points = 13), prior to calibration model development. During the PLS regression, two-thirds of spectra in one group were used for calibration equation development and the remaining one-third (every 3rd sample) spectra were used for model validation. To optimize the accuracy of PLS calibration models, different combinations of the spectral pretreatments were used. Leave-one-out cross-validation method was used, and the number of optimal factors selected for the PLS equation generally corresponded to the minimum of the predicted residual error sum of squares (PRESS). The saved regression equations were subsequently applied to the validation samples. Model accuracy and efficiency were assessed in the validation set on the basis of coefficient of determination (R^2), root mean square error of validation (SEP), and RPD. Usually, an optimal model should have higher R^2 and RPD and lower SEP.

RESULTS AND DISCUSSION

Relationship Between HVI_{area} and HVI_{count}. Figure 1 shows the relationship between HVI_{area} and HVI_{count} for a set of 406 samples. It suggests a strong correlation (Pearson correlation, R = 0.91), which is in good agreement with Farag's observation on a different data set in 2005 (R = 0.877). The scatter becomes more apparent as HVI_{area} and HVI_{count} (indicative of trash level) elevate, especially when HVI_{area} is greater than 0.40. If HVI_{count} was averaged for samples having the same HVI_{area} or HVI_{area} was averaged for samples having identical HVI_{count}, then resultant plots could be divided into two areas subjectively, with the boundary of HVI_{area} = 0.40 or HVI_{count} = 40 (Fig. 2). For example, they exhibit perfect linear relationships with smaller intercepts (closely passing the origin) and greater R (> 0.98) when $HVI_{area} \leq 0.40$ or $HVI_{count} \leq 40$ (or low trash cottons) than when $HVI_{area} > 0.40$ or $HVI_{count} > 40$ (or high trash cottons). Existence of more scattered samples and subsequent poorer correlation likely addresses the viewpoint that there might be larger dissimilarities between HVI_{area} and HVI_{count} in high trash cottons than in low trash ones. Nevertheless, it could describe the relationship between the two HVI trash presentations with either $HVI_{count} = 104.5*HVI_{area}$ or $HVI_{area} = 0.0093*HVI_{count}$. Obviously, these correlations are limited to low trash cottons.

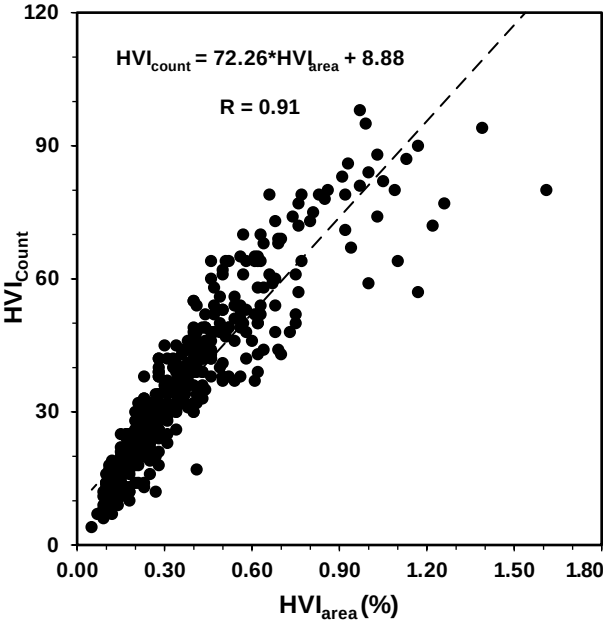

Figure 1. Relationship of HVI_{area} against HVI_{count} for a set of 406 samples.

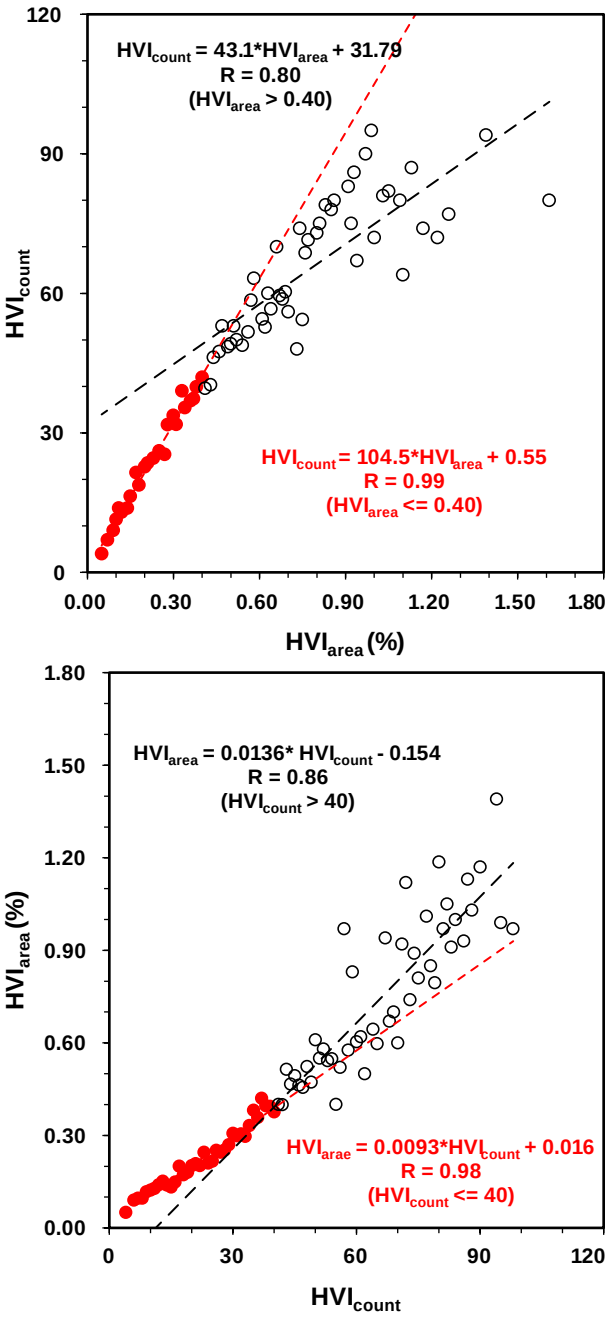

Figure 2. Relationship of HVI_{area} vs. averaged HVI_{count} (top) and HVI_{count} vs. averaged HVI_{area} (bottom).

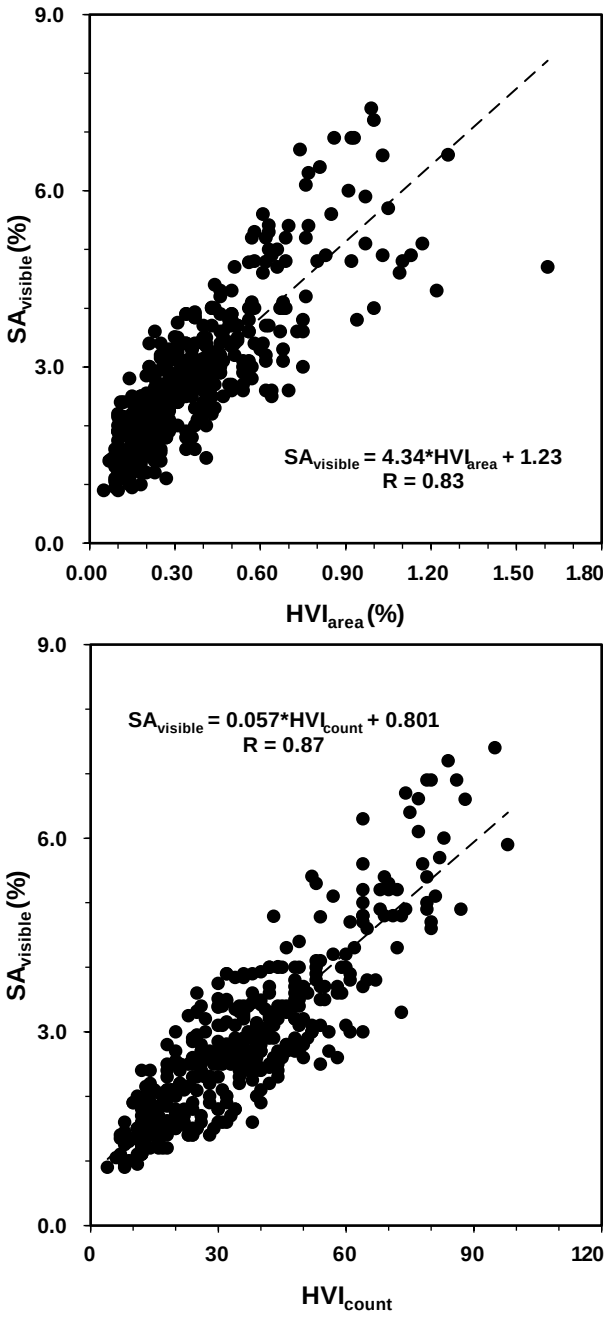

Figure 3. Relationship of HVI_{area} vs. $SA_{visible}$ (top) and HVI_{count} vs. $SA_{visible}$ (bottom).

Relationship Between HVI_{area} (or HVI_{count}) and $SA_{visible}$. Plots of both HVI_{area} vs. $SA_{visible}$ and HVI_{count} vs. $SA_{visible}$ (Fig. 3) reveal a strong correlation (R = 0.83-0.87), suggesting good agreement of the two testing methods in characterizing the cotton trash contents. The relationship between HVI_{area} and $SA_{visible}$ is close to that of HVI_{count} and $SA_{visible}$ (0.83 vs. 0.87).

If $SA_{visible}$ values were averaged for identical HVI_{area} or HVI_{count} (Fig. 4), they yield better correlations (R > 0.93) in low trash cottons than in high trash cottons, as anticipated. However, relationship between HVI trash and SA readings from Figs. 3 or 4 could not be conclusive because of (1) large intercepts in Fig. 3 and 4, (2) different sampling specimens between two trash determinations, and (3) the nature of heterogeneous and unexpected distribution of trash type and size within the sample.

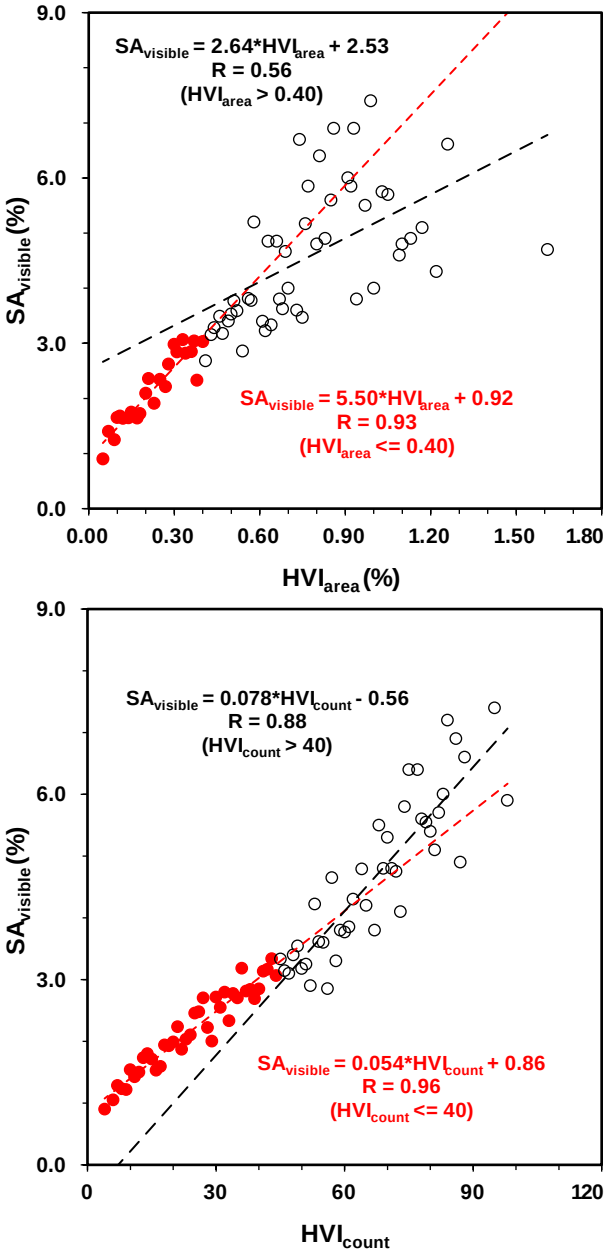

Figure 4. Relationship of HVI_{area} vs. averaged $SA_{visible}$ (top) and HVI_{count} vs. averaged $SA_{visible}$ (bottom).

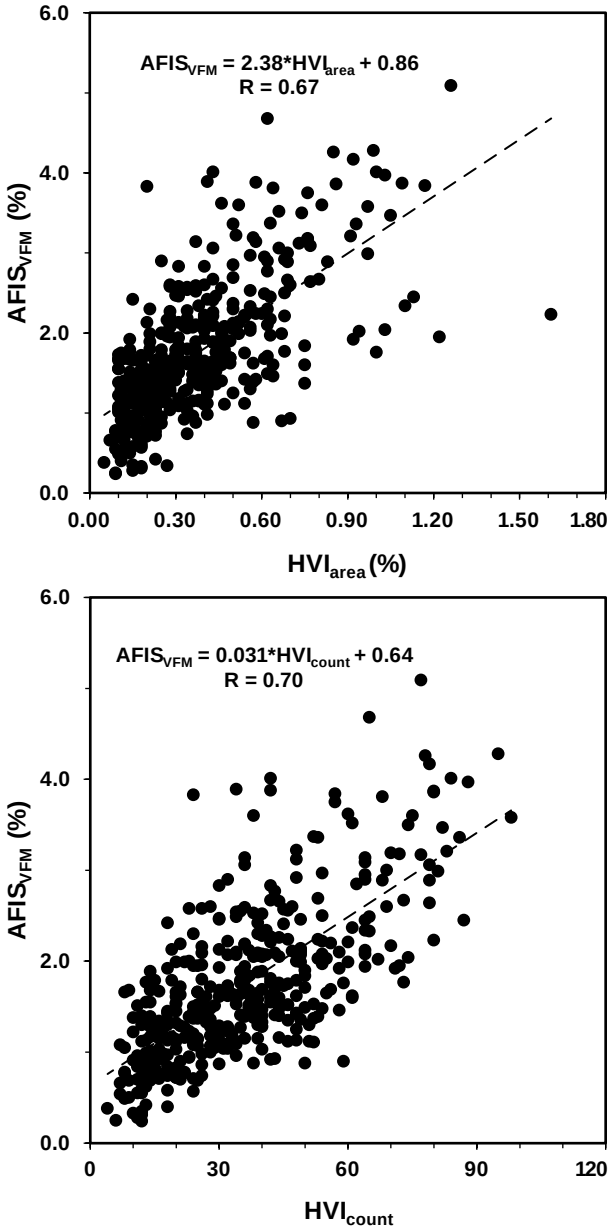

Figure 5. Relationship of HVI_{area} vs. $AFIS_{VFM}$ (top) and HVI_{count} vs. $AFIS_{VFM}$ (bottom).

Relationship Between HVI_{area} (or HVI_{count}) and $AFIS_{VFM}$. Although either HVI_{area} vs. $AFIS_{VFM}$ or HVI_{count} vs. $AFIS_{VFM}$ indicates a general trend in describing trash (Fig. 5), the correlations between the HVI and $AFIS_{VFM}$ (0.67-0.70) are much lower than those between the HVI and $SA_{visible}$ in Figure 3 (0.83-0.87). Similarly, both HVI_{area} and HVI_{count} against averaged $AFIS_{VFM}$ show a correlation of < 0.90 in low trash cottons (Fig. 6), which is slightly lower than that between HVI_{area} or HVI_{count} against mean $SA_{Visible}$ (Fig.4). It is expected, a major reason might be due to a large difference in sample amounts between HVI and AFIS measurement (10 g vs. 0.5 g). To this point, $SA_{visible}$ readings were further explored in this study.

Subjective Criteria to Determine the Relationship Between HVI_{area} (or HVI_{count}) and $SA_{visible}$. Going back to Fig. 3, the ratios between HVI_{area} and $SA_{visible}$ were calculated first. Relying on the magnitude of $HVI_{area}/SA_{visible}$ ratios, the samples were subjectively divided into five subsets (Fig. 7) and the resulting relationships from subsets are characterized in Table 2.

Table 2. Subjective criteria for subgruoping the samples with HVI_{area} / $SA_{visible}$ ratios.

$HVI_{area}/SA_{visible}$ ratios	Relationship	R
< 0.06 (Set A)	$SA_{visible} = 13.94*HVI_{area} + 0.50$	0.95
0.06 ~ 0.12 (Set B)	$SA_{visible} = 7.97*HVI_{area} + 0.53$	0.95
0.12 ~ 0.18 (Set C)	$SA_{visible} = 6.82*HVI_{area} + 0.09$	0.97
0.18 ~ 0.24 (Set D)	$SA_{visible} = 4.55*HVI_{area} + 0.28$	0.97
> 0.24 (Set E)	$SA_{visible} = 2.99*HVI_{area} + 0.69$	0.90

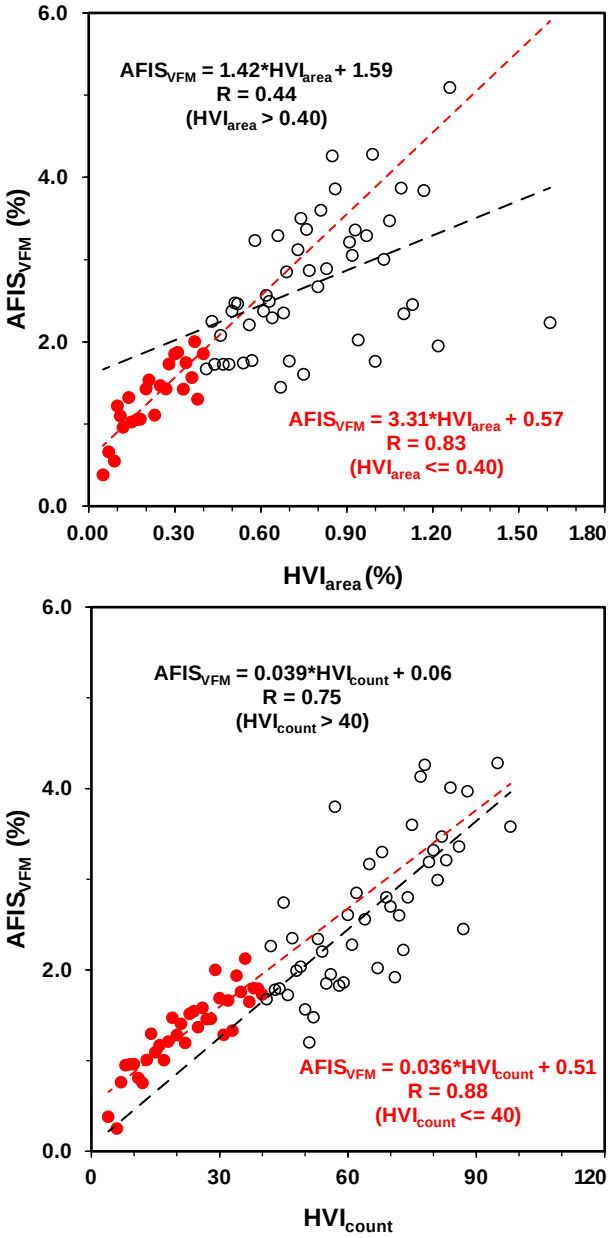

Figure 6. Relationship of HVI_{area} vs. averaged $AFIS_{VFM}$ (top) and HVI_{count} vs. averaged $AFIS_{VFM}$ (bottom).

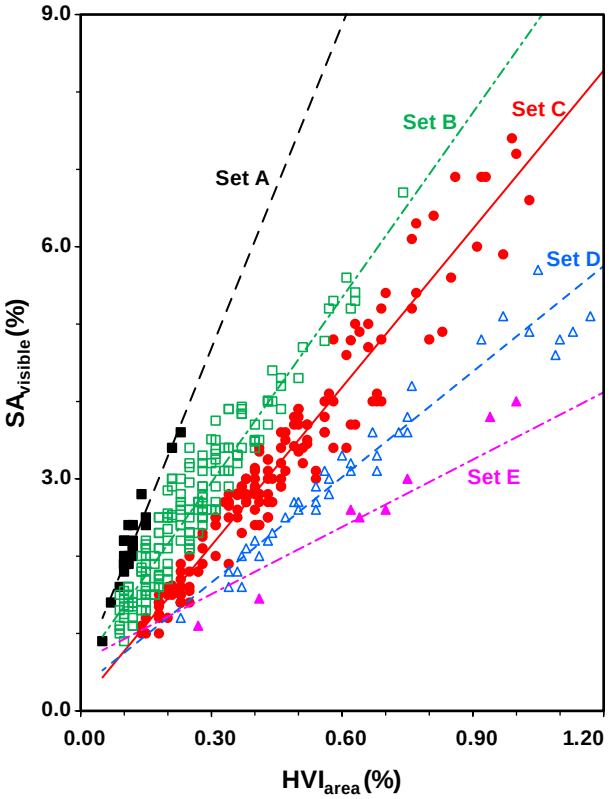

Figure 7. Plot of HVI_{area} against $SA_{visible}$ with 5 subjective subsets.

Table 2 suggests the higher correlation coefficients of > 0.90 for five small subsets, as predicted, and also varying intercepts ranging from 0.09 to 0.69. Of greatest interest is Set C that represents the $HVI_{area}/SA_{visible}$ ratio range of 0.12 to 0.18 and possesses the least intercept (0.09). Hence, the general conversion between HVI_{area} and $SA_{visible}$ can be expressed with the equation $SA_{visible} = 6.82*HVI_{area}$. Notably, the minimum intercept is desired in this approach, because not only both HVI and SA reading should be zero or close to zero if the sample does not contain any trash, but the two readings should change synchronously or appropriately if the existence of trash is uniform. Another factor is that the slope might represent the relationship between two trash values along horizontal and vertical axis.

Undoubtedly, the creation of five subsets is subject to more discussion. However, minimum consideration of this concept lies in both least subgroups and least intercept (i.e., one correlation line closely passing the origin). When trash exists homogeneously in samples, a plot of HVI trash against SA trash should yield a correlation line along the 45-degree direction. That is, samples should be found in an area adjacent to this correlation line (termed 45-degree area). In fact, occurrence of trash in samples is unpredictable, and HVI and SA analyze different portions in a sample. This

might produce a number of samples that are outside the 45-degree area, and those samples could be further separated into at least two clusters for either above or below the 45-degree area. Consequently, the use of five subsets in this study is appropriate and reasonable. Certainly, more work is necessary to better understand the effect of trash size, type, and weight distribution on HVI and SA measurements in the future.

If trash occurs uniformly and trash size/type are average, $HVI_{area}/SA_{visible}$ < 0.12 or > 0.18 might suggest the dominance of non-leaf or large-size trash and imply the heterogeneous presence of trash types and sizes that likely cause variations in trash contents between the two types of measurements. To interpret the practical samples, a margin of error of 5- to 10% or more should be considered.

By the same procedure, the samples were subjectively classified into five classes with the HVI_{count} / $SA_{visible}$ ratios (Table 3), and the representation of Fig. 3 is shown in Fig. 8. Similar to Table 2, Table 3 reveals the higher correlation coefficients of > 0.97 and different intercepts among the five subsets. The relationship between HVI_{count} and $SA_{visible}$ can be expressed by $SA_{visible} = 0.069*HVI_{count}$.

Table 3. Subjective criteria for classing the samples with HVI_{area} / $SA_{visible}$ ratios.

$HVI_{count}/SA_{visible}$ ratios	Relationship	R
< 6 (Set AA)	$SA_{visible} = 0.112*HVI_{count} + 0.496$	0.97
6 ~ 12 (Set BB)	$SA_{visible} = 0.085*HVI_{count} + 0.367$	0.98
12 ~ 18 (Set CC)	$SA_{visible} = 0.069*HVI_{count} + 0.106$	0.97
18 ~ 24(Set DD)	$SA_{visible} = 0.058*HVI_{count} - 0.143$	0.98
> 24 (Set EE)	$SA_{visible} = 0.047*HVI_{count} - 0.069$	0.99

From the respective equations of five subsets in Tables 2 and 3, the conversions between HVI_{count} and HVI_{area} are estimated to be 124.4, 93.8, 98.9, 78.4, and 63.6. The value of 98.9 is close to that 104.5 from the mean of HVI readings among low trash cottons, confirming the relationship between HVI_{count} and HVI_{area}.

NIR Model Verification of Relationship Between HVI_{area} and $SA_{visible}$. The relationship between HVI_{count} (or HVI_{area}) and $SA_{visible}$ can be validated by independent NIR spectral acquisition. This is because there are significant NIR spectral difference between lint fibers and trashes (Fortier et al., 2011; Liu et al., 2010b), leading to the hypothesis that appropriate trash reference should have the best NIR model performance. In this attempt, HVI_{area} readings were used as a reference to compare with the NIR spectral response.

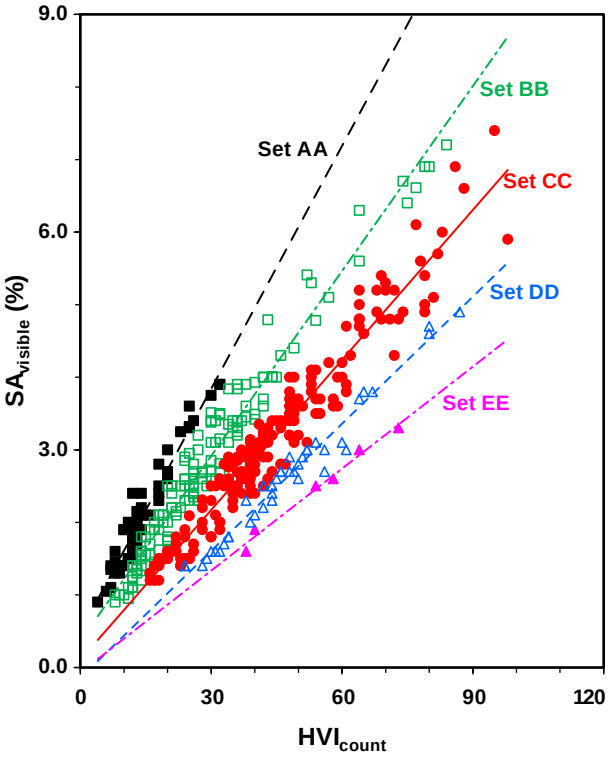

Figure 8. Plot of HVI_{count} against $SA_{visible}$ with 5 subjective subsets.

Sample Distribution and Calibration/Validation Assignment. With the HVI_{area} / $SA_{visible}$ ratios, all samples were subgrouped into respective classes (Table 4). Because of limited sample numbers, Sets A and E were not analyzed. For each of three other sets, on the order of the smallest to the largest in HVI_{area}, every third sample was selected to validate the calibration models that were built from the remaining samples.

Table 4. Sample distribution and assignment in each subset.

Subsets	Set A	Set B	Set C	Set D	Set E
Sample No.	20	183	152	40	11
Calibration No.		122	102	24	
Validation No.		61	50	16	

NIR Models on HVI_{area} Readings. Typical log (1/R) spectra in the 1100 to 2500 nm NIR region of three cotton fibers with the HVI_{area} readings of 0.18, 0.56, and 0.92 % are shown in Fig. 9. There are at least five intense broad bands (1490, 1935, 2105, 2270, and 2320 nm) mainly due to the 1st and 2nd overtones and combinations of OH and CH stretching vibrations of cotton cellulose (> 90% in total mass). Owing to relatively low trash amounts, spectral distinctions in this study are insignificant. However, there is one unique NIR band in the 2020 to

2200 nm region whose intensity decreases with total SA trash readings (Liu et al., 2010b) due to expected difference in color and compositions between trash and cotton fiber.

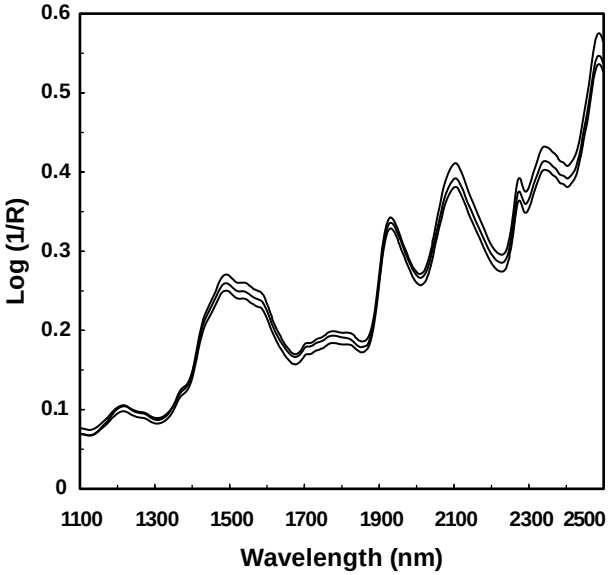

Figure 9. Typical NIR log (1/R) spectra of lint fibers with HVI_{area} (%) value of 0.18, 0.56, and 0.92

The statistics of optimal results from individual subsets are tabulated for comparison (Table 5). These models were obtained from the combination of mean center and Savitzky-Golay 1st derivative (2 degrees and 13 points) spectral preprocessing in the 1105 to 2495 nm region. To facilitate the comparison of NIR model performance between individual subsets, RPD and mean/SEP, quotients of SD and mean of reference values to SEP in validation sets, were used and also included in Table 5. In general, R^2, RPD, and mean/SEP from Set B are similar to those from Set C, and they are much better than ones from Set D.

Even though samples in Set B, C, and D were preselected on the basis of two isolated testings, that might not ensure homogeneous or close specimens for additional NIR measurement. Therefore, a 90% confidence interval was applied to exclude the outliers in three sets that had larger differences (or errors) between measured and NIR-predicted HVI_{area} from calibration and validation sets, respectively. The models were recalibrated and the results are also compiled in Table 5 (shown as bold italic). As expected, removal of outliers leads to the improvement of all model characteristics for each subset, and the difference among them might reflect different spectral response to the references in validation samples. Most notably, the model from Set C suggests the best model performance with the highest R^2, RPD, and mean/SEP. In other words, the best correlation might be related with the most appropriately determined references for HVI_{area} index. This observation is consistent with the simple statistical approach of least intercept (Table 2).

Comparative scatter plots of measured and NIR-predicted HVI_{area} from recreated models are shown in Fig. 10. It suggests how well the NIR model predictions agree with the references from a separate measurement. Set C exhibits a regression line more close to the 45-degree direction than Sets B and D.

A discussion of outliers and their detection is available in the literature (Boysworth and Booksh, 2001; Fearn, 2011). In our systematic studies on developing NIR models, we applied the 90% confidence interval rule to exclude outliers that showed large differences between measured and spectral model predicted properties (Liu et al., 2010a, 2010b, 2010c). In one of these reports (Liu et al., 2010c), we proposed

Table 5. Optimal NIR model statistics on HVI_{area}[z].

Subsets	Calibration set				Validation set					
	Range	SD	R^2	SEC	Range	SD	R^2	SEP	RPD	Mean/SEP
Set B	**0.09-0.74**	**0.125**	**0.83**	**0.051**	**0.09-0.63**	**0.127**	**0.74**	**0.064**	**1.98**	**4.0**
Set C	**0.14-1.03**	**0.207**	**0.91**	**0.063**	**0.15-0.99**	**0.198**	**0.75**	**0.102**	**1.94**	**4.3**
Set D	**0.23-1.26**	**0.297**	**0.91**	**0.088**	**0.36-1.13**	**0.205**	**0.45**	**0.195**	**1.05**	**3.2**
Set B[y]	***0.09-0.63***	***0.106***	***0.86***	***0.040***	***0.09-0.62***	***0.104***	***0.82***	***0.044***	***2.36***	***5.3***
Set C[y]	***0.14-1.03***	***0.197***	***0.95***	***0.045***	***0.15-0.92***	***0.186***	***0.90***	***0.062***	***3.00***	***6.7***
Set D[y]	***0.23-1.26***	***0.296***	***0.96***	***0.058***	***0.37-1.13***	***0.224***	***0.63***	***0.139***	***1.62***	***4.8***

[z] SD, standard deviation; SEC, root mean square error of calibration; SEP, root mean square error of prediction; RPD, ratio of SD to SEP. Mean/SEP, ratio of mean to SEP.

[y] By applying 90% confidence interval, samples with greater differences between NIR predicted and actual HVI_{area} values were excluded from the calibration and validation sets.

a prescreening procedure to determine appropriate calibration samples on the basis that two strength readings from two independent strength tests should be highly correlated. However, even though samples were prescreened, that might not ensure the use of homogeneous or close specimens for additional NIR spectral collection, hence a 90% confidence interval was applied to exclude the outliers.

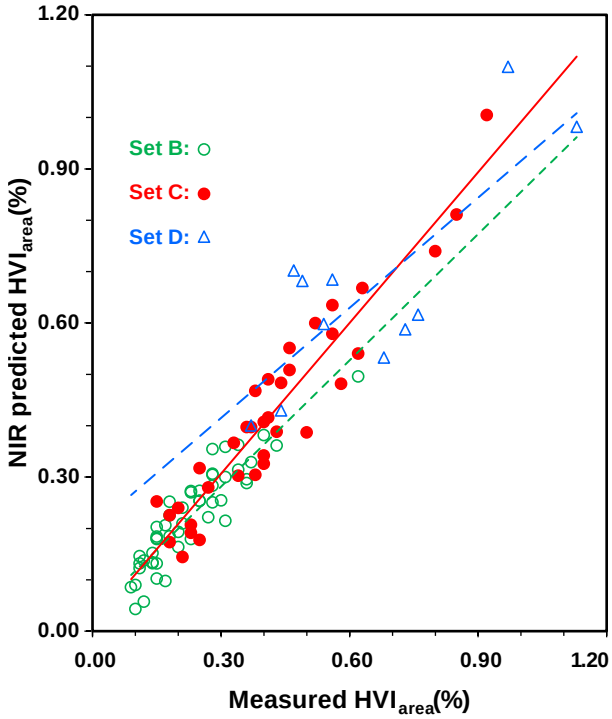

Figure 10. Plot of measured vs. NIR predicted HVI_{area} for Set B (○ and dashed), Set C (● and solid), and Set D (△ and dashed).

The number of outlier samples depends on the selected level of confidence interval, and the application of a 90% confidence interval in these studies is subjective and therefore debatable. It is likely that not every sample outside the 90% confidence interval is a true outlier. Thus, the use of outlier samples to represent the ones with large differences (or errors) between measured and predicted properties might be inappropriate, instead, these samples could be considered as not meeting the required (or expected) quality.

CONCLUSIONS

Although a strong correlation exists between two HVI trash readings from identical specimens, their mean values yield a relationship of HVI_{count} = 104.5*HVI_{area} in low trash cottons ($HVI_{area} \leq 0.40$). With regard to the relationship between HVI trash readings and those obtained by either SA or AFIS, it is reasonable to have scatter plots of HVI_{area} (or HVI_{count}) against $SA_{visible}$ and $AFIS_{VFM}$, despite similar trends in describing trash contents. Overall, the correlations between HVI_{area} (or HVI_{count}) and $SA_{visible}$ are stronger than those between HVI_{area} (or HVI_{count}) and $AFIS_{VFM}$.

Regardless of apparent challenges from such factors as trash uniformity, type, size, and measuring methods, the samples were subgrouped subjectively according to the ratios of $HVI_{area}/SA_{visible}$ (or $HVI_{count}/SA_{visible}$). From the respective plot of individual subsets, two general conversions, $SA_{visible}$ = 6.82*HVI_{area} and $SA_{visible}$ = 0.069*HVI_{count}, are best representative. These general conversions were derived from the subset where the $HVI_{area}/SA_{visible}$ ratio range was between 0.12 to 0.18 or where the $HVI_{count}/SA_{visible}$ ratio range was between 12 to 18. Undoubtedly, these two conversion constants might change with relative amount of trash size and type and also their weight distribution.

To verify the proposed conversion between HVI_{area} and $SA_{visible}$, NIR spectra were correlated with HVI_{area} readings in three subsets. Considering the heterogeneous distribution of trash in fibers and different sampling specimens between NIR spectral and HVI reference measurement, a 90% confidence interval was applied to exclude outlier samples from the calibration and validation sets. The redeveloped models exhibit different response to various references of validation samples in three subsets. Of interest is that the model from Set C (i.e., the $HVI_{area}/SA_{visible}$ ratio range of 0.12-0.18) suggests the highest R^2, RPD, and means/SEP, likely demonstrating better determined references for the samples in this subset than those in other two sets (Set B and D).

ACKNOWLEDGMENTS

We are grateful to Mr. James Knowlton (USDA, AMS) for his encouragement during this study. Also, we appreciated the technical assistance from D. Martha, N. Carroll, F. Patricia, and M. Morris (ARS, Clemson) for collecting the reference and spectral data.

DISCLAIMER

Mention of a product or specific equipment does not constitute a guarantee or warranty by the USDA and does not imply its approval to the exclusion of other products that might also be suitable.

REFERENCES

American Society for Testing and Materials (ASTM). 2007a. Standard test method for non-lint content of cotton (D2812). In Annual Book of ASTM Standard Standards. Vol. 07.01. ASTM, West Conshohocken, PA.

American Society for Testing and Materials (ASTM). 2007b. Standard test method for neps in cotton fibers (AFIS-N instrument) (D5866). In Annual Book of ASTM Standards. Vol. 07.01. ASTM, West Conshohocken, PA.

American Society for Testing and Materials (ASTM). 2007c. Standard test method for measurement of physical properties of cotton fibers by high volume instruments (D5867). In Annual Book of ASTM Standards. Vol. 07.01. ASTM, West Conshohocken, PA.

Anthony, W.S. 2007. The harvesting and ginning of cotton. p. 176–202. *In* S. Gordon and Y.-L. Hsieh (eds.) Cotton: Science and Technology. Woodhead Publishing Limited, Cambridge, England.

Boysworth, M.K., and K.S. Booksh. 2001. Aspects of multivariate calibration applied to near infrared spectroscopy. p. 209–240. *In* Burns, D.A. and E.W. Ciurczak (eds.) Handbook of Near-Infrared Analysis. Marcel Dekker Inc., New York, NY.

Farag, R. 2005. Evaluation of key hypotheses associated with HVI and AFIS fiber quality measures. p. 2258-2270 *In* Proc. Beltwide Cotton Conf., New Orleans, LA. 4-7 Jan. 2005. Natl. Cotton Counc. Am., Memphis, TN.

Fearn, T. 2011. Improving accuracy by rejecting outliers? NIR News 22 (6):12 -14.

Fortier, C.A., J.E. Rodgers, M.S. Cintron, X. Cui, and J.A. Foulk. 2011. Identification of cotton and cotton trash components by Fourier-transform near-infrared spectroscopy. Textile Res. J 81:230 - 238.

Funk, P.A., C.B. Armijo, A.T. Hanson, Z.A. Samani, M.A. Macias-Corral, G.B. Smith, and J.T. Riordan. 2005. Converting gin and dairy wastes to methane. Trans. ASAE 48:1197–1201.

Gordon, S. 2007. Cotton fiber quality. p. 68–100. *In* S. Gordon and Y.-L. Hsieh (eds.) Cotton: Science and Technology. Woodhead Publishing Limited, Cambridge, England.

Haanstra, W.G., W. Hansen, M.J.G. Huys, B.J. Kip, P. Palmen, J. Roumen, M. Snieder, T. van de WeerdhoF, S. Wiedemann, and V.A.L. Wortel. 1998. Information depth of vis-NIR light in polyethylene films using transmission and reflectance measurements. Appl. Spectroscopy 52:863–868.

Knowlton, J.L. 2002. Development in cotton standards. *In* Proc. Beltwide Cotton Conf., Atlanta, GA. 8-13 Jan. 2002. Natl. Cotton Counc. Am., Memphis, TN.

Liu, Y., G. Gamble, and D. Thibodeaux. 2010a. Evaluation of 3 cotton trash measurements methods by visible/near-infrared reflectance spectroscopy. ASABE paper no. 1008708.

Liu, Y., G. Gamble, and D. Thibodeaux. 2010b. Assessment of cotton fiber and trash contents in lint cotton waste by UV/visible/near-infrared reflectance spectroscopy. J. Near Infrared Spectroscopy 18:239–246.

Liu, Y., G. Gamble, and D. Thibodeaux. 2010c. An approach to determine calibration samples in NIR modeling for cotton fiber strengths. NIR News 21: 9–11,13.

Montalvo Jr., J.G., and G.J. Mangialardi. 1983. Systems errors in Shirley Analyzer measurements. Textile Res. J. 53: 408–414.

Thomasson, J.A., and S.A. Shearer. 1995. Correlation of NIR data with cotton quality characteristics. Trans. ASAE 38:1005–1010.

USDA, AMS. 2005. Cotton Classification, Understanding the Data. USDA AMS report, April, 2005. Available t: http://www.ams.usda.gov/AMSv1.0/getfile?dDocName=stelprdc5074569 (verified 21 Sept. 2012).

Wakelyn, P.J., J.V. Edwards, N.R. Bertoniere, B.A. Triplett, L. Hunter, A.D. French, M.-A. Rousselle, D.D. McAlister, D.P. Thibodeaux, W.R. Goynes, and G.R. Gamble. 2007. Cotton Fiber Chemistry and Technology. CRC Press, Boca Raton, FL.

Xu, B., C. Fang, R. Huang, and M.D. Watson. 1997. Chromatic image analysis for cotton trash and color measurements. Textile Res. J. 67:881–890.