Chapter 12

Sticky Cotton Sampling

S.E. Naranjo and E.F. Hequet

Sampling is a fundamental component of any research program and is an essential element for accurately measuring and quantifying the characteristics of cotton lint quality for both research and commercial grading purposes. A sample is a set of "sample units" that allows one to make inferences about the entire population from which these observations are drawn. Sampling activities are guided by a structured set of rules called a sampling plan or program. The sampling plan includes the designation of the sample unit, how sample units are spatially allocated among potential sample units in the population, and how many sample units will be collected for each sample in order to get a reliable mean estimate.

In sampling for lint stickiness there may be different goals depending on the stage at which observations are made (for example, field, gin, or textile mill). Crop monitoring during the season and use of decision-making tools to aid in determining the need for sweetpotato whitefly (or cotton aphid) control to prevent sticky cotton development (chapter 7) or the use of remedial actions to reduce or eliminate stickiness (chapter 10) could potentially allow growers to produce high quality lint and avoid price penalties. Estimation of stickiness in harvested cotton is an obvious consideration for the textile manufacturer to prevent costly machinery downtime and excessive machinery maintenance. At what stage or stages in the crop production lint stickiness should be determined remains an open question. Overall, the most critical issue for cotton producers and textile manufacturers is that, wherever the sticky cotton determination is made, it accurately predict possible textile processing problems.

Stickiness Measurement Systems

Sampling for lint stickiness is a two-stage process: (1) collection of sample units from the field, module, or bale and (2) assay of these sample units to provide a quantitative measurement of stickiness. There are a number of different measurement systems that have been developed to qualitatively or quantitatively assess lint stickiness (chapter 13; see also Hector and Hodkinson 1989).

The sampling methods and plans that will be described here are limited to three physically based measurement systems; however, the approaches and analyses would be similar regardless of measurement methodology. These systems include the manual sticky cotton thermodetector (SCT) which is currently the method recommended by the International Textile Manufacturer's Federation for measuring cotton stickiness (Perkins and Brushwood 1995), the high speed stickiness detector (H2SD) (Hequet et al. 1997), and the fiber contamination tester (FCT) (Mor 1996). The SCT involves spreading a thin web of conditioned lint between aluminum foil sheets, heating under pressure, separating the aluminum foil sheets, and counting the number of adhering sticky spots (Brushwood and Perkins 1993). The H2SD, with a few minor modifications, essentially duplicates the process of the SCT on an automated basis, greatly speeding sample throughput. For the FCT a fiber sliver, whose mass and length is fixed, is fed into a microcard. The web that is formed passes between two heated drums under pressure. The sticky spots adhering to the drum are counted with an image analyzer. More detail is provided in chapter 13.

The measurement instrument employed is a significant factor in the development of any sampling plan. For example, a plan developed for the SCT is not directly applicable to the H2SD because each platform has its own inherent error characteristics and variability.

Sampling for Cotton Lint Stickiness

There has been considerable research and development of methods and machinery for the measurement of lint stickiness. However, very little research has addressed the basic issues of sampling and the development of sampling plans for the accurate estimation of stickiness. In this chapter we will provide a detailed summary of our current knowledge of sampling for lint stickiness at both preharvest and postharvest stages in the cotton production and processing cycle.

The authors appreciate the funding by Cotton Incorporated, which partially defrayed the cost of this research.

Data Sources and General Methods

The data in support of preharvest field sampling work was collected from central Arizona and southern California between 1995 and 1999. Over this period, data was collected using a variety of different sample units from a total of 87 field sites, some of which included samples from the same field on several different dates. In most cases we also determined the amount of time necessary to collect each sample unit for further analyses of sampling efficiency. Seed cotton from these field sample units was ginned using a small research gin, and following hand blending a subsample was then assayed using the SCT or the H2SD, or both, depending on the amount of lint available and the underlying objectives of the project each year. All H2SD assays were completed by the International Textile Center at Texas Tech University in Lubbock, TX; most SCT assays were done at the USDA-ARS laboratory in Clemson, SC. In general, three replicate assays were conducted on both instruments.

Several studies have been conducted at the International Textile Center to evaluate postharvest methods for estimating cotton stickiness in bales and modules. In one experiment, 50 Texas bales representing a range of stickiness were selected. Ten 1-pound sample units were taken per bale, and for each sample unit three replicate assays were conducted on FCT, H2SD, and SCT instruments at several locations. A second experiment was conducted following the same sampling protocol on 100 bales coming from California and Arizona that were selected to represent a large range of stickiness. To further define within-module variability, a third study was undertaken consisting of 283 modules from Arizona and California. For each module a single sample unit (similar to the grader's sample, about one-half pound of lint) was taken from three bales for each module. Three replicate assays were conducted for each sample unit on the H2SD.

Comparison of Sample Units

Proper selection of the sample unit can reduce bias by ensuring that the unit is representative of the universe (a field, for example) being sampled. Further, selecting the sample unit that minimizes both variance and cost can optimize the efficiency of sampling. All of the sample units we evaluated for field sampling here are representative of the sample universe, but differed in the level and extent of aggregation (table 1). Whole-plant sample units are unbiased because they encompass all of the lint on a single or multiple plants that represent quantifiable units of the entire field. Boll sample units also are unbiased because the individual bolls in any one sample unit (for example, 20 or 40 bolls) are selected at random within the crop canopy and again represent a quantifiable unit of the habitat. In bale or module sampling the goal was to develop sampling protocols that are compatible with current grader sampling methods. Thus, the sample unit was not the subject of further experimental work and consisted of at least 4 ounces of lint taken from each side of the bale.

For field sampling we generally found that regardless of the size of the sample unit, ranging from lint collected from 20 open bolls at random (1 boll per plant) to all of the lint on 30 consecutive plants (table 1), mean estimates of stickiness were essentially the same using either the SCT or H2SD platforms. However, from the perspective of sampling efficiency, the best sample unit is the one that provides the highest level of precision or repeatability for the lowest cost. Larger sample units sometimes had comparatively lower variance, but they were more time-consuming, and thus more costly, to collect from the field. Southwood (1978) suggested that the relative net precision (RNP) of a sample unit should be proportional to $1/(C_u S_u)$, where C_u is the cost per capita of the sample unit and S_u is a measure of sample unit's relative variability. Higher values of RNP indicate a more efficient sample unit (better precision at a lower cost). Here we use the coefficient of variation (CV = SD/mean) to represent relative variation and sample collection time in the field to estimate costs per unit. Based on results averaged over 5 years, the 1-plant sample unit was most efficient, followed by the 20-boll sample unit for both assay platforms. This tells us that smaller sampling units are more efficient than larger units. Further discussion on field sampling will focus only on the 1-plant and 20-boll sample units.

Sampling Distributions

We contrasted the sampling distributions of thermodetector counts from the SCT and the H2SD for field samples and from the SCT, H2SD, and the FCT for bale and module samples. We calculated the coefficient of dispersion (CD), estimated as the ratio between the sample variance and sample mean, to characterize the between-assay and between-sample unit sampling distributions. Generally,

CD < 1 indicates a regular distribution, CD ≈ 1 indicates a random or Poisson distribution, and CD > 1 indicates an aggregated or clumped distribution. For field samples we found that CDs between replicate assays (within-assay) indicated a more regular to random distribution for the SCT, but an aggregated distribution for the H2SD (table 2). Likewise, CDs for between-sample unit counts were lower for SCT than the H2SD and again indicated a random distribution for the SCT and an aggregated distribution for the H2SD (table 2).

Based on studies conducted on Sudanese commercial cotton bales, Fonteneau Tamine et al. (2000) found that the CD of stickiness readings using the H2SD was approximately 4.84, leading the authors to reject the hypothesis of a Poisson distribution. The authors fitted the data to an empirical model that relates the mean to the variance. This empirical model indicated an aggregated distribution.

Based on sampling studies of bales from Texas, it appears that all within-assay CD's are well above 1, revealing an aggregated distribution (table 3). The mean CD values were close to 2 for the H2SD and SCT instruments tested. Assays on the FCT revealed an even more aggregated distribution, suggesting inconsistent results on this instrument. Except for the FCT, the within-bale CD's are all around 1. Thus, the variability within a bale is smaller than the variability within a sample unit. Similar conclusions can be drawn from the second set of bale samples from Arizona and California (table 4). Consequently, for U.S. cottons, it appears that the regular classer's sample unit should be representative of the entire bale. A final study to evaluate within-module variability showed that the within-assay CDs averaged 1.5, revealing a slight overdispersion relative to a Poisson distribution (table 5). The within-module CD averaged 2.8, revealing that the variance of stickiness readings within a module is roughly twice the variance within a bale. This indicates that classification of stickiness based on module averaging is not feasible because of the large degree of variability in stickiness within a module.

Module averaging consists of testing each bale, averaging all of the bales from the same module, then applying this average value to each individual bale. In doing this we could incur the risk of overestimating or underestimating the stickiness value of the individual bales. This may have an extremely negative effect on both producers and spinners. The overestimated bales will be discounted with no reason and the underestimated bales will lead to cotton mixes with a higher than desired stickiness level. Consequently, we cannot envisage module averaging for stickiness based on a single grader's sample unit per bale; thus, each bale should be tested individually.

Partitioning of Variance Components

Thermodetector assays are conducted on lint from a field or bale sample unit. Thus, there are two sources of variation: (1) variability among replicate assays from individual sample units and (2) variability among sample units collected from the same field or bale. Because the sampler can exert some level of control over both of these sources of variation, we quantified and evaluated their contributions to overall sample variation. Nested ANOVA was used to partition and quantify within- and between-sample unit variability for a set of samples assayed on the SCT and the H2SD. For the SCT we found that approximately 57 percent of the variation was attributable to differences among field sample units (field variation) while the remaining 43 percent represented between-assay variability (laboratory variation). This latter source of variation for the SCT includes variability caused by subsampling and the SCT operator. Because the H2SD largely eliminates operator error we would expect the laboratory component of variation to decline. Instead we found that nearly 70 percent of the total variance was attributable to between-assay error for the H2SD, while only 30 percent was attributable to field variation. The probable cause for this result will be explained below under "Other Sources of Variation."

This variance partitioning analysis can be used to determine the optimal allocation of sampling effort between the field and laboratory components (Cochran 1977, Southwood 1978) as—

$$N_L = (C_F S_L / C_L S_F)^{0.5}$$

where:

N_L is the laboratory sample size,
C_F is the cost per unit of field sampling,
C_L is the cost per unit lab assay,
S_F is the field variance, and
S_L is the laboratory variance.

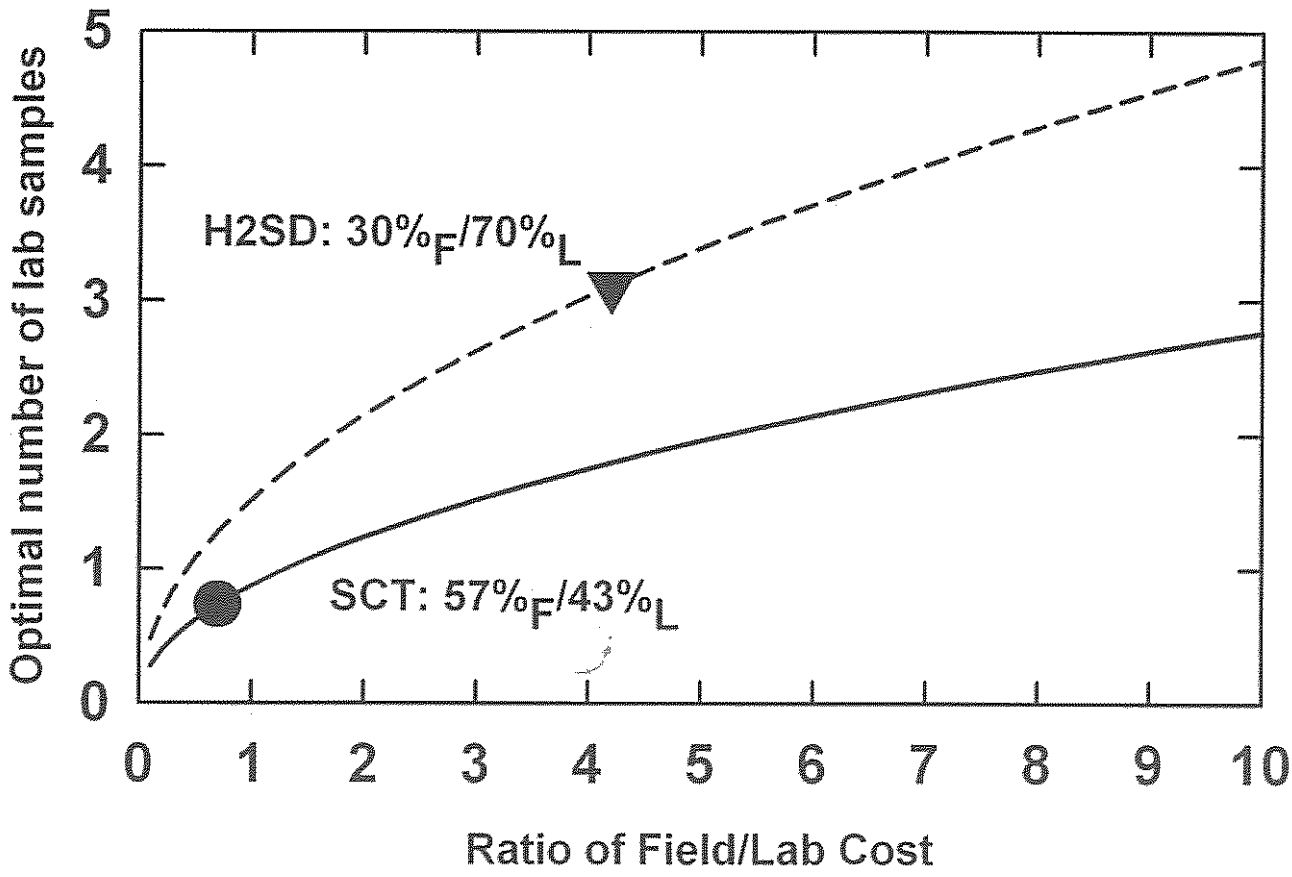

Figure 1. Components of variation and implications for sampling effort allocation relative to costs. The symbols denote the optimal number of laboratory assays from each field sample unit based on known field collection costs and the assumption that single H2SD and SCT assays take 0.5 and 3 min, respectively, to complete. Results based on 20-boll sample units collected during 1998-99.

Using this approach we calculated the optimal number of replicate assays necessary to minimize variance in relation to cost (figure 1). Assuming a field cost of about 2 minutes per unit (for a 20-boll sample unit) and an SCT assay cost of 3 minutes, this analysis suggests that only a single assay should be conducted on each sample unit (figure 1, circle on the solid line). Assuming a field cost of 2 minutes per unit and an H2SD assay cost of 0.5 minutes, our analysis suggests that 3 assays should be conducted on each sample unit (triangle on dotted line). These analytical results follow directly from the more qualitative patterns shown in table 2 and simple cost considerations. The more regular distribution of counts between assays for the SCT and the high cost of assay suggest that sampling effort is better spent on the collection of the more variable field sample units rather than replicate assays on each unit. The reverse is essentially true for the H2SD, for which sampling distributions are aggregated for both assay and field, but assay costs are much lower than field collection costs. Interestingly, our results for the H2SD agree with standard assay protocols already in place for the SCT and H2SD, which call for 3 replicate assays for each sample unit.

Other Sources of Variation

There are additional sources of variation that can influence the estimation of lint stickiness. Two of these are worthy of further discussion here: variation between laboratories conducting the assays, and the degree of preparation of lint samples prior to assay. During field studies conducted in 1996, samples (20-boll sample unit) collected from 18 different field sites were subsequently assayed on SCT platforms run by two different laboratories. A nested ANOVA of

square-root transformed counts was used to estimate the variance component due to differences between laboratories within the context of within- and between-sample-unit variability as described above. Results pooled over all field sites indicated that only about 9 percent of the total variation was attributable to differences between laboratories. However, despite this relatively small amount of variability, further analyses demonstrated significant differences in mean spot counts among the 18 fields (figure 2). Further, there was a consistent pattern in the difference between the two laboratories, with higher spot counts being reported by one laboratory and differences being greater with increased average spot counts. In the absence of a standardized test methodology, and more importantly of standards to calibrate the instruments, this type of difference is to be expected.

USDA uses a very careful procedure to pick the cotton used for HVI (high volume instrument) calibration. The following is an extract from "The Classification of Cotton" (USDA/AMS 1993): "As a first step the USDA conducts an intensive search for the most uniform bales of cotton in the current crop. Candidate bales are screened for uniformity of fiber quality by testing 12 samples drawn from throughout each bale. Bales that pass this preliminary screening then undergo detailed analysis to determine whether they meet USDA standards for certification and use as calibration cottons." None of this exists for stickiness, and thus it makes between-laboratory comparisons extremely difficult. Still, it does emphasize the potential importance of human and perhaps machine error in the assay process. Based solely on sampling distributions, it would appear that there is relatively little variation between different laboratories using the SCT and only minor variation with the H2SD for stickiness estimation in bale sampling (see tables 3 and 4).

Another source of variation concerns the degree of lint cleaning before a sample is assayed. This is unlikely to be an important factor for commercially ginned seed cotton, or for classing samples, but small research gins used by researchers generally have no capacity for cleaning lint to efficiently remove seed fragments, leaf trash, and other debris. We evaluated the effects

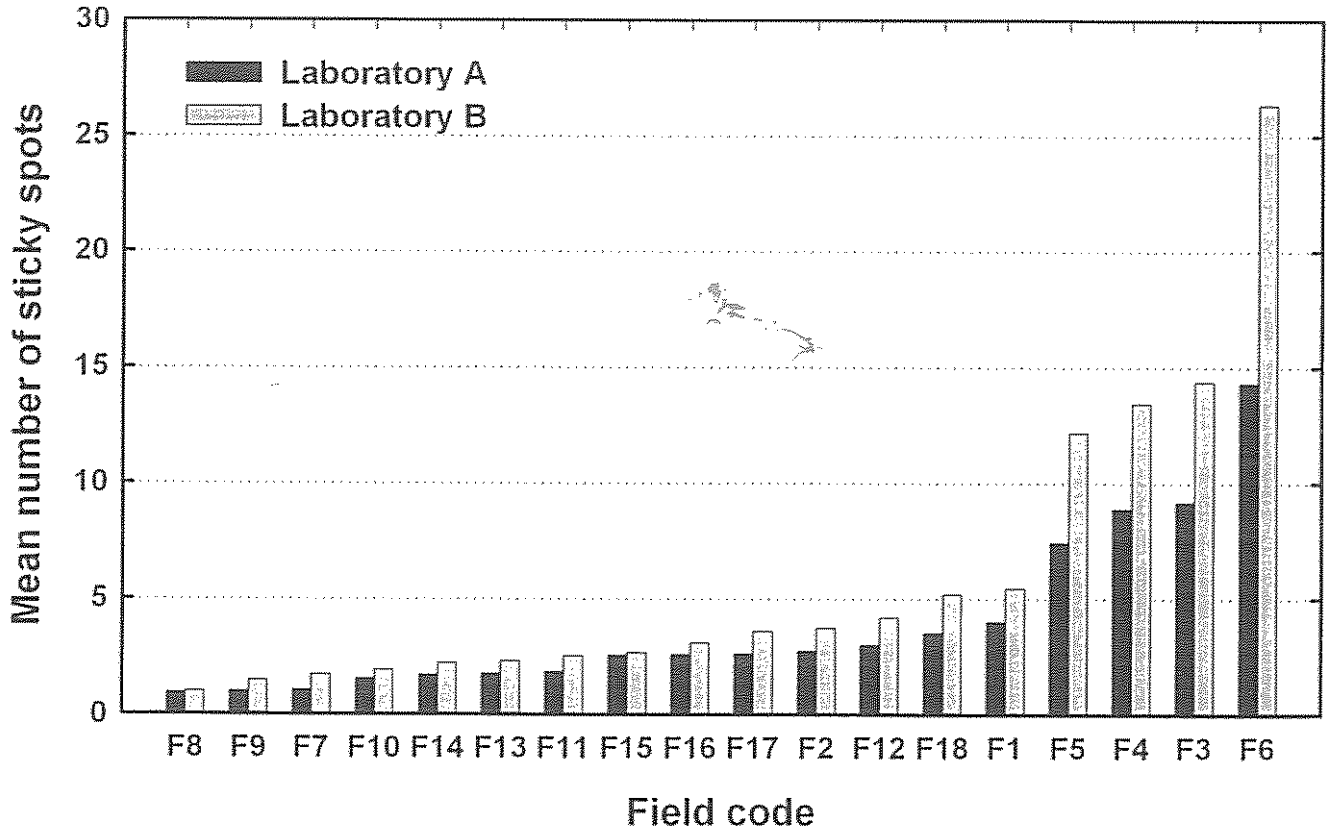

Figure 2. Comparison of stickiness estimates for 18 fields between two laboratories using the SCT instrument.

of lint cleaning on spot counts and variability for 32 samples collected from the field in 1998 and 1999. Prior to assay on both the SCT and the H2SD, ginned (research gin) samples were either subjected to further cleaning or not. Samples to be cleaned were processed through a single pass of the Shirley Analyser. Three replicate assays were performed on each instrument. We found essentially no change in sampling characteristics from the SCT as a result of lint cleaning (table 6). On average, there was no significant change in the magnitude of the SCT reading and measures of sampling distributions (CD) changed very little. Conversely, lint cleaning had a dramatic effect on the sampling characteristics from the H2SD. Cleaning reduced the average spot count by more than 30 percent, and the between-assay sampling distribution changed from aggregated to random (table 6). Cleaning did not significantly alter the size distribution of spots ($\chi^2 = 1.86$, $P = 0.39$); overall, 66, 15, and 17 percent of the spots were categorized as small, medium, and large, respectively. The automated platform would appear to be more sensitive to lint trash and other impurities, which results in higher and more variable counts between assays. The lack of change with cleaning in the SCT suggests that some trash is removed from the samples during sample preparation and that the technician is able to more readily differentiate spots caused by sticky lint rather from those caused lint trash. For the H2SD it suggests that sampling properties can be greatly altered by lint cleaning and that consistent protocols need to be followed in comparative research studies where cleaning capacity may or may not be available.

Sampling Plans for SCT and H2SD Platforms

The sampling plan is a procedure for collecting a sample from the field or from a module or bale and arriving at an estimate of stickiness with a desired level of repeatability and accuracy. This includes the sample unit (in the case of field sampling) used to take samples, the number of sample units to collect, the processing of the seed cotton, the assay platform, and the number of replicate assays to perform. Here we assume that 3 replicate assays will be performed on each sample unit.

The two remaining elements required to complete a sampling plan for estimating lint stickiness are the interrelated factors of sample size and precision. For this we use an empirical model that allows the estimation of the sample variance based on the sample mean. We used Iwao's patchiness regression (Iwao 1968) to describe the relationship between the sample mean and variance. Sample size curves were estimated for two levels of precision (figure 3) using the general relationship:

$$N = (t_{\alpha/2}/D)^2(s^2/m^2),$$

where:

N is sample size,
t is Student's t for a specified α (type I error rate),
D is a fixed proportion of the true mean,
m is the sample mean, and
s^2 is sample variance estimated from the empirical model.

The cost of sampling was estimated by multiplying the sample size by the per unit costs of sample collection and assay as discussed above. For bale sampling we estimated the sample collection costs at 1 minute per unit. Specific examples of sample size and cost are presented in table 7 for four levels of precision. For example, given a true mean of 10 sticky spots on the SCT, a sample size of 23-25 would permit us to estimate a mean between 9 and 11 with 95 percent confidence. A sample size of 58 would be required to make the same estimate from bales samples using the H2SD. Regardless of precision, it can be seen that sample size requirements decline as levels of stickiness increase. The level of precision desired also influences sample size requirements dramatically, with lower levels of precision requiring many fewer sample units. For field sampling, the speed of the H2SD results in generally lower sampling costs despite higher sample size requirements, especially at lower levels of precision. The 20-boll sample unit appears to be more cost-efficient for field sampling. Overall, the sample size chosen by the user will depend on an interplay between cost considerations and how much precision and accuracy is required in determining levels of stickiness.

Another way to examine the relationship between sample size and precision is to ask, "What sample size would be needed to confidently distinguish levels of stickiness between two cottons?" This is a critical issue if stickiness is to become a standard measure of lint quality. We can address this question by calculating statistical power (1-β) where β is the Type II error, the probability of accepting a false null hypothesis. The greater the power, the greater is our ability to confidently distinguish between two alternatives.

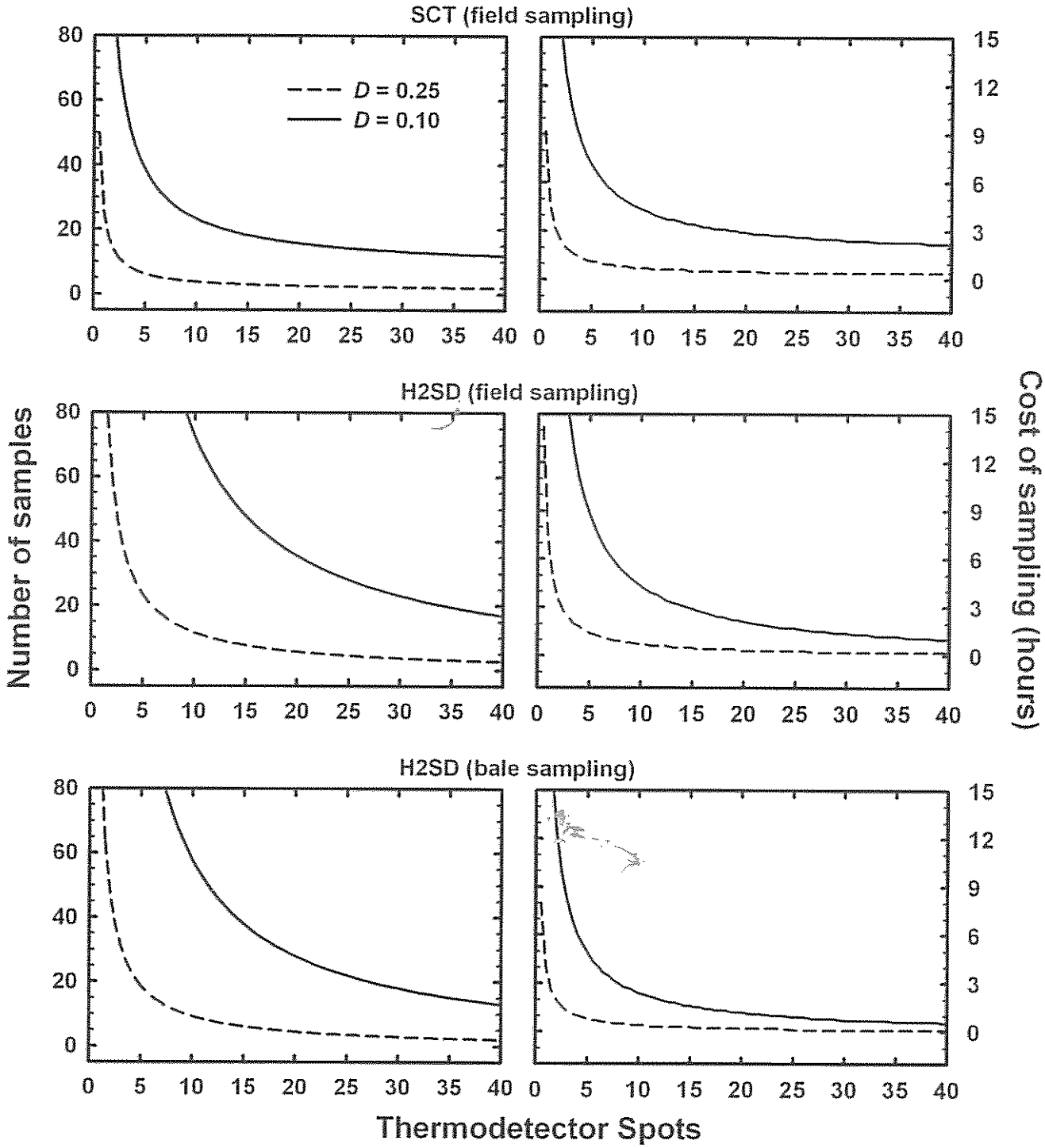

Figure 3. Sample sizes and sampling costs required to estimate mean stickiness with two levels of precision (expressed as a fixed proportion of the mean with 95% confidence) on two assay instruments and assuming a mean-variance relationship described by Iwao's patchiness regression. Field sample sizes are based on a 20-boll sample unit. Bales samples are based on a grader's sample unit. See text for details.

To determine power we need to define α, the Type I error rate, and δ, the numerical difference we wish to distinguish. Here, we set $\alpha = 0.05$ and let δ vary. We also need to square-root-transform the data to normalize the intrasample variance.

We will assume here that a spot count of around 10 represents the division between nonsticky and sticky lint and examine sampling properties associated with this level of stickiness. In figure 4 we plot statistical power for field sampling as a function of sample size for different levels of δ, the spot count difference that we can expect to detect when the true mean is 10 on the SCT. With a $\delta = 0.5$ on a square-root scale, a relatively large sample size would be required for an adequate level of power (> 90 percent). For example, a sample size of about 16 would be needed to discriminate two samples with arithmetic sticky counts of 7.1 and 13.4 with a power of 90 percent. Low sample sizes can provide high power but only at low levels of resolution. For instance, a sample size of 5 would be expected to discriminate two samples with counts of 1.4 and 26.6 spots, on an arithmetic scale, with 90 percent power. As noted above, the sample size chosen by the user will depend on cost considerations and the goals of the sampling program. For example, relatively low power and resolution may be adequate for a researcher interested in distinguishing between two alternative control methods for suppressing sweetpotato whiteflies in the field. Alternatively, very high power and resolution will be needed for determining stickiness as part of lint quality assessment. This latter issue is explored in more detail below.

From the bale and module sampling results presented above ("Sampling Distribution"), it is clear that the FCT exhibits a higher level of variability than both the SCT and the H2SD. Furthermore, even if the variability of the SCT is acceptable, from a practical point of view, its usage is, and will remain, extremely limited (because of, for example, operator effect or use of a manual instrument). Consequently, the H2SD appears to be a better candidate for large-scale testing of lint quality. Here we calculate power curves for the H2SD using data from the 100 bales from Arizona and California discussed above. Figure 5 shows that to discriminate between 2 samples with 90 percent power, 6 replications would be necessary with $\delta = 1$, 4 replications with $\delta = 2$, and 3 replications with $\delta = 3$.

Another way to examine this problem is to define two categories, sticky and nonsticky. For the purpose of this demonstration we assume that the threshold between sticky and nonsticky is 10. Let's further assume that we have a bale with a H2SD spot count of six. Is this count statistically below 10 sticky spots? The answer to this question with an average sample size of 6 is yes with a power of 80 percent (figure 6). A sample size of 9 from the same bale with 6 sticky spots is statistically below 10 spots with a power of 90 percent (figure 6).

Relationship Between Preharvest and Postharvest Stickiness

Although preharvest sampling may serve several goals as discussed previously, it is of interest to understand the relationship between these field estimates of stickiness and those determined from harvested cotton. Over the course of the field studies described here we often collected subsamples of lint after machine harvesting. Typically these were collected immediately or within 1 day after collecting the final in-field samples. Linear regression analysis was used to describe the relationship between the stickiness of final field samples and harvest samples after square root transformations of both counts. For the SCT the regression model is given as $[\textit{In-field} \times 0.79] + 1.01$ ($r^2 = 0.55$, $n = 53$), indicating that harvest stickiness was consistently higher than stickiness measured from samples drawn from the field, at least for mean stickiness levels less than 25 on an arithmetic scale. The mean difference in counts was about 0.5 on a square-root scale. For the H2SD the regression model is given as $[\textit{In-field} \times 1.01] + 2.96$ ($r^2 = 0.71$, $n = 24$) indicating that harvest stickiness was consistently higher than stickiness determined from field samples at all levels of stickiness. Here, the mean difference in spot counts was about 3 on a square-root scale. The reason for these consistent differences is unclear but could be related the spreading and mixing of honeydew droplets or possibly the addition of insect sugars from stems and leaf parts during the harvest process. In any case, our results suggest that field sampling may slightly underestimate stickiness of the harvested lint and more research may be needed to evaluate this issue.

Sequential and Classification Sampling

Sampling plans for the estimation of lint stickiness or the categorization of stickiness could potentially be made more efficient through the use of what are known as sequential sampling plans. Such plans are commonly used in entomology for both research purposes and for decision-making in integrated pest management

(Binns et al. 2000). Both efficiency and precision are optimized because in sequential sampling the need for further sample information is assessed following the collection of each individual sample unit. For the estimation of a mean, the method ensures that no more sample units are collected than necessary in order to achieve a predetermined level of precision. Further, because in most cases sample size decreases with increasing means (figure 4) the method automatically ensures that the correct number of sample units are collected without any prior knowledge of the mean. Sequential sampling for mean estimation operates by accumulating counts (in this case sticky spots) over subsequent sample units and then consulting a cumulative count curve or a table to determine the need for more sample units or the termination of sampling. Once sampling is terminated, the mean is calculated by simply dividing the cumulative count by the number of sample units collected.

Operationally, the method would be most simple for single-stage sampling in which the count can be made immediately after collecting the sample unit (for example, counting whitefly adults on cotton leaves). Because stickiness sampling is a two-stage process requiring laboratory assay, sampling for stickiness in the field would require that the user collect a set number of sample units. However, a sequential plan could then be implemented at the assay stage, ensuring that only as many sample units as necessary be processed to meet a predetermined precision. With the SCT, which requires approximately 3 minutes to complete a single assay, substantial time could potentially be saved. Less time would be saved using a faster platform (such as the H2SD), but sequential sampling may still prove a valuable cost-saving approach. Cost saving could be even more significant for bale testing because the first stage of sampling

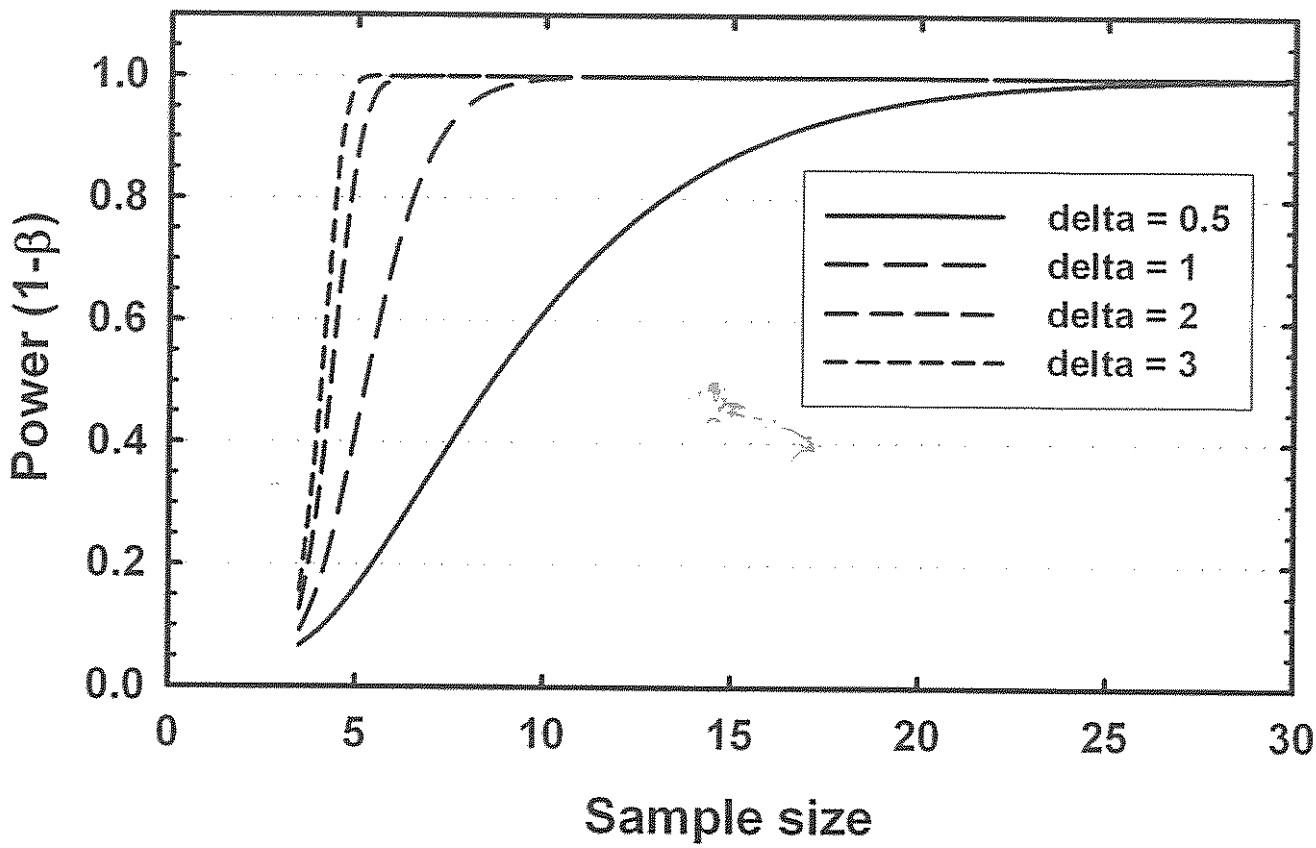

Figure 4. Statistical power for square-root transformed SCT spot counts from field samples as a function of sample size for various levels of delta, the difference we can expect to detect when the true mean is 10 on an arithmetic scale (2-tailed [$\mu = \mu_0$] with $\alpha = 0.05$ and $\sigma^2 = 0.241$).

(lint collection) is much quicker and in some instances the assay machinery may be near the lint source (gin, textile mill, classing office, etc.).

A second application of sequential sampling involves the classification of lint stickiness rather than estimation of mean stickiness levels per se. In the case where one simply wants to determine whether lint is sticky or nonsticky, a sequential classification approach could save substantial time and effort. Operationally, the method is similar to that described above for mean estimation except that a critical density must be specified. In insect control this critical density would be the economic threshold. Densities above the critical density would require control; densities below would require no action. For sticky cotton it would be the level of stickiness delineating two classes of stickiness, be it the difference between nonsticky and lightly sticky or the difference between lightly and moderately sticky.

To demonstrate this approach, let's assume that 10 sticky spots is the critical threshold. Several approaches have been described for sequential classification sampling. Here we will apply the method developed by Wald known as the sequential probability ratio test (Binns et al. 2000). For field sampling using the SCT we will assume a Poisson sampling distribution. For bale sampling using the H2SD we used the Taylor Power Law ($s^2 = am^b$) to estimate the relationship between the mean (m) and variance (s^2). We further set Type I and II error rates to 0.05, set the minimum sample size to 1, and use the simulation methods of Binns et al. (2000) to evaluate the sampling plan.

Results are shown in figure 7 for three different maximum sample sizes (3, 10, and 25) on each instrument. The average sample number simply shows the sample size that would be required to classify lint stickiness as a function of the level of stickiness. What is immediately clear is that very few sample units are

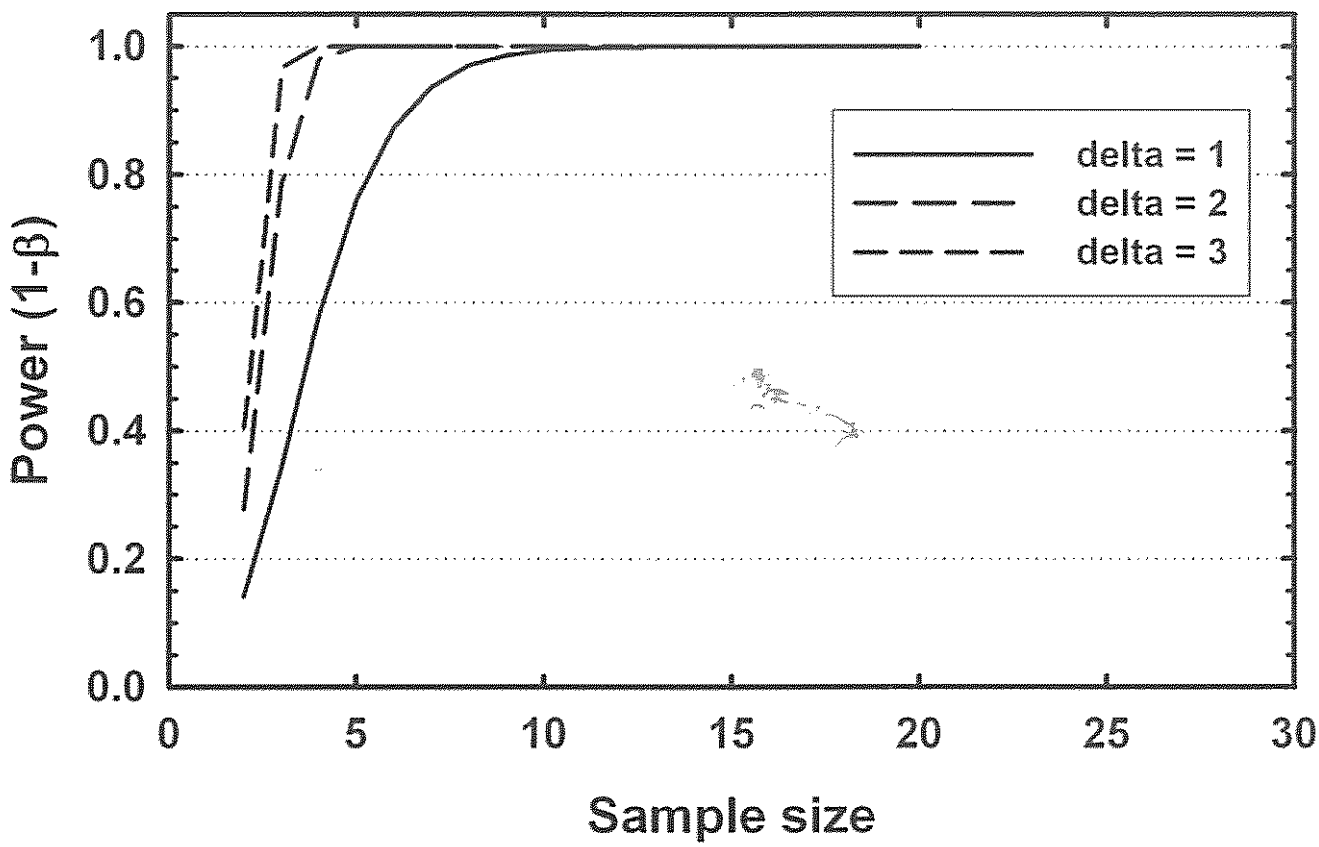

Figure 5. Statistical power for square-root transformed H2SD spot counts from bale samples as a function of sample size for various levels of delta, the difference we can expect to detect when the true mean is 10 on an arithmetic scale (2-tailed [$\mu = \mu_0$] with $\alpha = 0.05$ and $\sigma^2 = 0.392$).

required when stickiness is below or above 10 spots, but that a relatively large sample size is required when stickiness is at or near this critical density. The steepness and width of the sample size curve about the critical density depends on the maximum sample size. The operating characteristic shows the probability of classifying the lint as nonsticky as a function of the number of sticky spots. Ideally, this curve would be vertical at the critical density resulting in perfect discrimination between sticky and nonsticky cotton. In reality, there is the possibility of misclassification, with greater error associated with lower maximum sample sizes and changes in other sampling parameters (not shown). For example, with a maximum sample size of 25, readings below about 9.6 and above about 10.5 would be classified correctly as nonsticky or sticky, respectively. These boundaries widen as maximum sample size declines. Samples with a mean stickiness very near 10 will be misclassified roughly 50 percent of the time. The Poisson distribution of the field SCT data results in narrower sample size functions compared with the aggregated distributions of the bale H2SD data. However, there is relatively little difference in the error curves.

This approach demonstrates that sampling efficiency could be improved dramatically by ensuring that maximal effort is expended only when lint stickiness is near the critical density. Classification of lint stickiness outside this narrow range would require very little effort. Again, the implementation of such a plan would depend on the goals and purpose of sampling. Such an approach may still be unfeasible for classing purposes under current sampling protocols, but it may be useful for research purposes where more time and effort can be devoted to sampling.

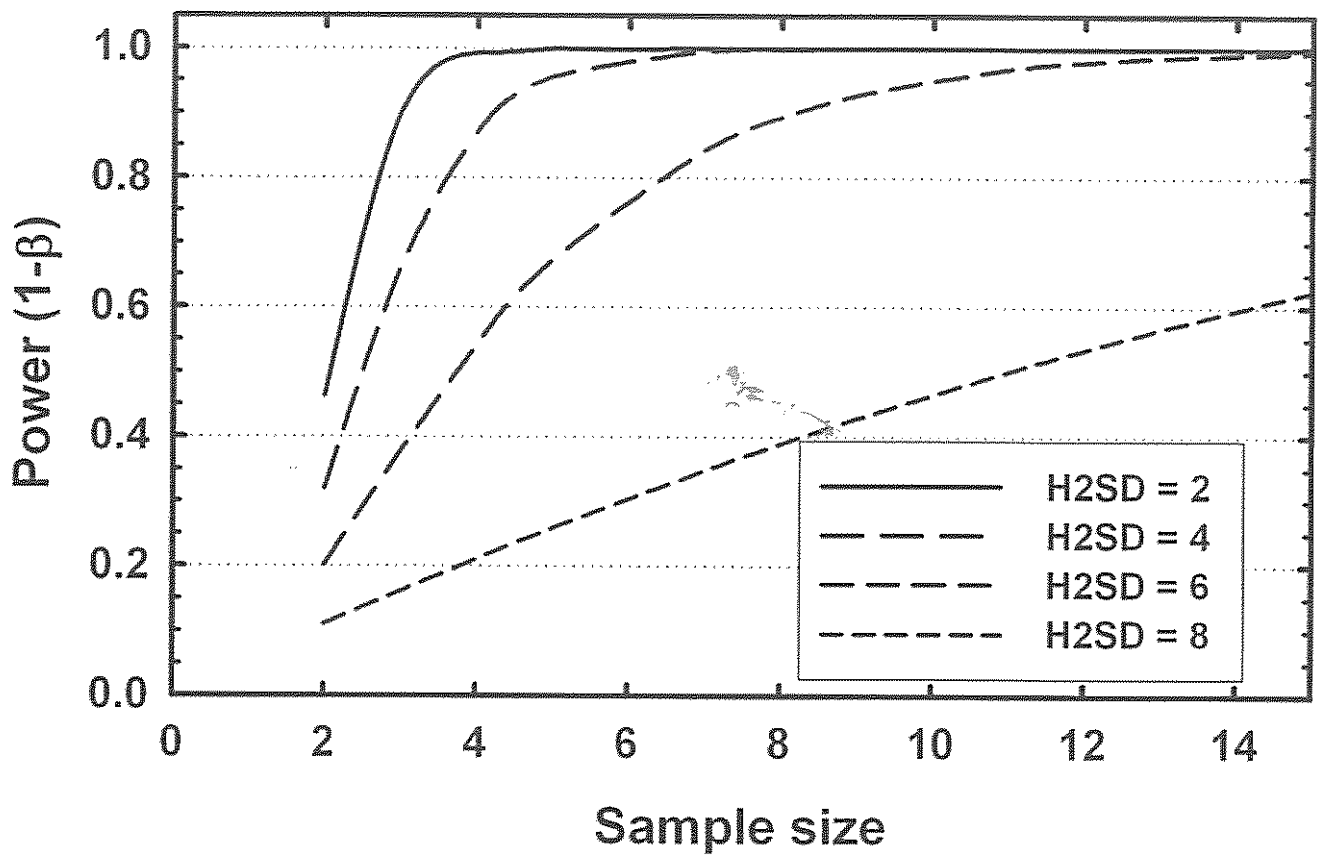

Figure 6. Statistical power for square-root transformed H2SD spot counts from bale samples as a function of sample size for various differences in stickiness from a true mean of 10 on an arithmetic scale (1-tailed [$\mu \geq \mu_o$] with $\alpha = 0.05$ and $\sigma^2 = 0.392$).

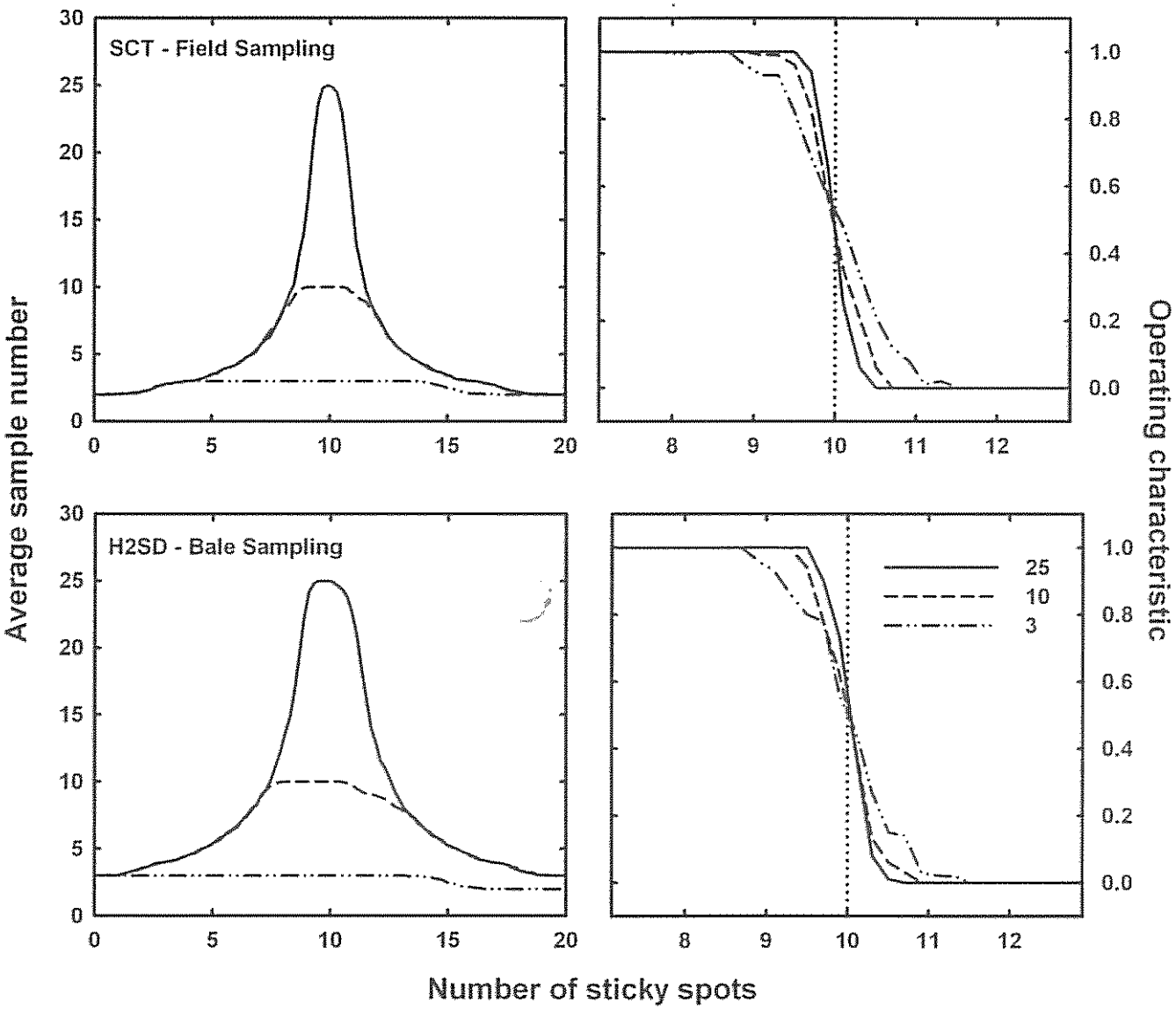

Figure 7. Performance of various sequential sampling plans for classifying lint stickiness from field and bale samples assayed on SCT and H2SD instruments. This example assumes that 10 sticky spots is the critical boundary between sticky and nonsticky lint. The average sample number simply shows the required sample size as a function of stickiness levels. Sample size requirements are maximal near the critical boundary. The operating characteristic gives the probability of classifying the lint as nonsticky as a function of stickiness levels. An ideal curve would be vertical at 10 as denoted by the dotted line. The lines represent a maximum sample size of 3, 10, or 25.

Conclusions and Future Developments

Here we have presented a detailed summary of our current knowledge of sampling for lint stickiness at both preharvest and postharvest stages of production and suggest sampling protocols for two measurement instruments. Relative to a grader's sample, the precise estimation of lint stickiness using either the SCT or the H2SD requires considerably more effort regardless of whether the determination is made directly from within-field samples or from harvested modules or bales. The FCT has shown even greater variability, and we did not pursue the development of sampling plans for this instrument.

Depending on the level of precision desired and the level of lint stickiness, as many as 23 sample units or as few as 2 sample units may be required to estimate stickiness of field-collected lint on the SCT. An equivalent range for the H2SD is 73 to 3 sample units for field sampling and 58 to 2 sample units for bale sampling. While these sampling requirements may be feasible for research purposes in the field, gin, or textile mill, they are clearly unsuitable for lint quality assessment in a commercial setting. It is simply not feasible with current thermodetection technology to delineate sticky from nonsticky cotton with acceptable power and precision using a grader's bale sampling method, which consists of a single sample unit that is assayed in triplicate. The use of module averaging, in which stickiness is determined from sampling of multiple bales per module and then assigning that level of stickiness to each bale is also unfeasible. Our results show that variability between bales in a module is even larger than within-bale variability. This system could incur the risk of overestimating or underestimating the stickiness value of the individual bales.

It is also very unlikely that producers could effectively use in-field sampling for assessing the dynamics of lint stickiness in their production systems. Even though sample size requirements for field sampling of stickiness are modest compared to requirements for most insect pest sampling (for example, see Naranjo 1996), the time necessary to collect an adequate sample are relatively large and currently there is the additional constraint of limited access to testing machinery. The application of a sequential sampling protocol could enhance the efficiency of sampling for research purposes, but even the simple classification of stickiness would require more effort than grader's sampling for sufficient confidence in the classification outcome.

Although we did not explore the sampling properties of other stickiness testing methods, including chemically based tests, our overall conclusions regarding commercial feasibility are unlikely to change. Regardless of whether samples are collected from the field or from modules or bales, stickiness of ginned lint is simply too variable to achieve reasonable precision with only a single or a few sample units, especially when stickiness levels are between 0 and 10. One practical alternative would be to develop an online assay system that could accommodate the throughput necessary to test a larger number of sample units. The current mechanically based systems discussed here would be impractical for this purpose, but perhaps some type of spectroscopic measurement (or example, Fourier-transform infrared spectroscopy or Raman spectroscopy [chapter 13]) would be rapid enough. Whether or not such systems can be used to estimate stickiness is currently unknown; however, there are several laboratories in the United States examining the potential of Fourier-transform infrared spectroscopic analysis for this purpose.

Table 1. Summary of field sample units examined 1995-1999 using manual (SCT) and automated high-speed (H2SD) sticky cotton thermodetectors

		SCT		H2SD	
Sample unit	Cost per unit[a]	RNP[b]	*n*	RNP[b]	*n*
	minutes				
1 Plant	1.40	0.031	29	0.033	6
2 Plants	2.18	0.016	35	–	–
5 Plants	4.51	0.010	22	–	–
10 Plants	8.40	0.005	22	–	–
20 Plants	16.17	0.003	22	–	–
30 Plants	23.95	0.002	17	–	–
20 Bolls	2.08	0.018	55	0.026	38
40 Bolls	3.54	0.015	26	0.019	21
50 Bolls	4.27	0.010	15	0.013	10
80 Bolls	6.45	0.017	5	–	–
100 Bolls	7.91	0.011	5	–	–
200 Bolls	15.20	0.005	5	–	–

[a] Cost in time to collect sample unit in the field.

[b] Relative net precision (1/[(SD/mean)×cost]) measures the relationship between relative variability and cost; higher values indicate a more efficient sample unit.

Table 2. Coefficients of dispersion (variance/mean) for 3 replicate assays on each instrument and for sample units collected from the same field

[All samples collected 1998-1999 using 20-boll and 1-plant sample units]

		Coefficient of dispersion			
	Instrument	Median	Maximum	Minimum	*n*
Within assay	SCT	0.62	8.24	0.00	320
	H2SD	2.53	84.79	0.00	320
Between samples	SCT	1.04	4.30	0.08	44
	H2SD	3.24	9.96	0.46	44

Table 3. Coefficients of dispersion (variance/mean) calculated on 50 bales from Texas for three types of instruments (SCT, FCT and H2SD) at three laboratories

[10 sample units per bale with 3 replicate assays per sample unit]

	Instrument	Laboratory	Coefficient of dispersion		
			Mean	Minimum	Maximum
Within assay	FCT	A	5.8	0.0	77.3
	H2SD	B	2.4	0.0	21.5
	SCT	B	2.6	0.0	23.6
	SCT	C	2.6	0.0	36.6
Within bale	FCT	A	2.2	0.8	8.7
	H2SD	B	0.6	0.1	1.4
	SCT	B	0.7	0.1	1.4
	SCT	C	1.3	0.1	8.7

Table 4. Coefficients of dispersion (variance/mean) calculated on 100 bales from Arizona and California for three types of instruments (FCT, H2SD and SCT) at two laboratories

[10 sample units per bale with 3 replicate assays per sample unit]

	Instrument	Laboratory	Coefficient of dispersion		
			Mean	Minimum	Maximum
Within assay	FCT	D	6.7	0.0	541.2
	H2SD	D	2.7	0.0	35.6
	H2SD	B	2.1	0.0	14.3
	SCT	B	2.7	0.0	33.9
Within bale	FCT	D	4.4	0.3	112.4
	H2SD	D	2.7	0.1	30.1
	H2SD	B	0.9	0.1	4.4
	SCT	B	1.7	0.1	9.0

Table 5. Coefficients of dispersion (variance/mean) calculated on 283 modules from Arizona and California for one type of instrument (H2SD) at one laboratory

[1 sample unit from each of 3 bales per module with 3 replicate assays per sample unit]

	Instrument	Laboratory	Coefficient of dispersion		
			Mean	Minimum	Maximum
Within assay	H2SD	D	1.5	0.0	11.9
Within module	H2SD	D	2.8	0.0	67.7

Table 6. Comparison of stickiness estimates and variability on cleaned and raw lint assayed by SCT and H2SD

	SCT		H2SD	
	Raw	Clean	Raw	Clean
Range[a]	1.3-14.3	0.3-18.0	3.3-71.0	0.7-65.0
Median CD[b]	0.48	0.50	2.01	0.87
Mean % change from raw	–	-1.41	31.61	–
t[c]	–	0.22	9.22	–
P	–	0.83	<0.01	–

[a] Range of spot counts (untransformed) for the 32 samples.

[b] CD: coefficient of dispersion (variance/mean) for replicate assays.

[c] *t*-test to evaluate the null hypothesis that the mean % change in spot count = 0

Table 7. Sample sizes required to estimate various levels of mean stickiness with various levels of precision using different sample units on two assay platforms.

[Results expressed as a fixed proportion of the mean with 95% confidence and assume a mean-variance relationship described by Iwao's (1968) patchiness regression. Numbers in parentheses are estimates of the total time (hours) required to complete sampling, including both field or bale and laboratory]

	Mean sticky spot count			
Precision	**10**	**20**	**30**	**40**
SCT (1-plant unit)				
0.1	25 (4.4)	17 (2.8)	13 (2.3)	12 (2.0)
0.15	11 (2.0)	7 (1.3)	6 (1.0)	5 (0.9)
0.2	6 (1.1)	4 (0.7)	3 (0.6)	3 (0.5)
0.25	4 (0.7)	3 (0.5)	2 (0.4)	2 (0.3)
SCT (20-boll unit)				
0.1	23 (4.3)	16 (2.9)	13 (2.4)	12 (2.2)
0.15	10 (1.9)	7 (1.3)	6 (1.1)	5 (1.0)
0.2	6 (1.1)	4 (0.7)	3 (0.6)	3 (0.5)
0.25	4 (0.7)	3 (0.5)	2 (0.4)	2 (0.4)
H2SD (1-plant unit)				
0.1	155 (7.5)	73 (3.5)	45 (2.2)	31 (1.5)
0.15	69 (3.3)	32 (1.6)	20 (1.0)	14 (0.7)
0.2	39 (1.9)	18 (0.9)	11 (0.5)	8 (0.4)
0.25	25 (1.2)	12 (0.6)	7 (0.3)	5 (0.2)
H2SD (20-boll unit)				
0.1	73 (4.3)	35 (2.1)	23 (1.4)	17 (1.0)
0.15	32 (1.9)	16 (0.9)	10 (0.6)	7 (0.4)
0.2	18 (1.1)	9 (0.5)	6 (0.3)	4 (0.3)
0.25	12 (0.7)	6 (0.3)	4 (0.2)	3 (0.2)
H2SD (bales)				
0.1	58 (2.4)	28 (1.2)	18 (0.7)	13 (0.5)
0.15	26 (1.1)	12 (0.5)	8 (0.3)	6 (0.2)
0.2	15 (0.6)	7 (0.3)	5 (0.2)	3 (0.1)
0.25	9 (0.4)	5 (0.2)	3 (0.1)	2 (0.1)

References

Binns, M.R., J.P. Nyrop, and W. van der Werf. 2000. Sampling and Monitoring in Crop Protection. CAB International, Wallingford, UK.

Brushwood, D.E., and H.H. Perkins, Jr. 1993. Cotton stickiness potential as determined by minicard, thermodetector, and chemical methods. *In* D.J. Herber and D.A. Richter, eds., Proceedings Beltwide Cotton Conferences, January 10-14, 1993, New Orleans, LA, pp. 1132-1135. National Cotton Council of America, Memphis, TN.

Cochran, W.G. 1977. Sampling Techniques, 3rd ed. John Wiley and Sons, New York

Fonteneau Tamine, O., E. Goze, and R. Frydrych. 2000. Qualitative classification of cotton stickiness in the H2SD high speed stickiness detector. Textile Research Journal 70:866-871.

Hector, D.J., and I.D. Hodkinson. 1989. Stickiness in cotton. ICAC Review Articles in Cotton Production Research No. 2. CAB International, Wallingford, UK.

Hequet, E., R. Frydrych, and M. Watson. 1997. The use of the high speed stickiness detector on a large range of cotton coming from different countries. *In* P. Dugger and D.A. Richter, eds., Proceedings Beltwide Cotton Conferences, January 6-10, 1997, New Orleans, LA, pp. 1649-1653. National Cotton Council of America, Memphis, TN.

Iwao, S. 1968. A new regression method for analyzing the aggregation pattern of animal populations. Researches on Population Ecology (Kyoto) 10:1-20.

Mor, U. 1996. FCT—Fiber Contamination Tester—a new instrument for the rapid measurement of stickiness, neps, seed-coat fragments and trash—for the ginner to the spinner. *In* P. Dugger and D.A. Richter, eds., Proceedings Beltwide Cotton Conferences, January 9-12, 1996, Nashville, TN, pp. 205-212. National Cotton Council of America, Memphis, TN.

Naranjo, S.E. 1996. Sampling *Bemisia* for research and pest management applications. *In* D. Gerling and R.T. Mayer, eds., *Bemisia:* 1995 Taxonomy, Biology, Damage, Control and Management, pp. 209-224. Intercept, Andover, UK.

Perkins, H.H., Jr., and D.E. Brushwood. 1995. Interlaboratory evaluation of the thermodetector cotton stickiness test method. *In* D.A. Richter and J. Armour, eds., Proceedings Beltwide Cotton Conferences, January 4-7, 1995, San Antonio, TX, pp. 1189-1191. National Cotton Council of America, Memphis, TN.

Southwood, T.R.E. 1978. Ecological Methods, 2nd ed. Chapman and Hall, London.

USDA/AMS [United States Department of Agriculture, Agricultural Marketing Service]. 1993. The classification of cotton. Handbook 566.