Yield Assessment in COMPACI in Eastern and Southern African Countries: History, Evolution, and The Current Approach

11 DECEMBER 2013

PRESENTED TO:
DEG (Deutsche Investitions- und Entwicklungsgesellschaft mbH)
Kämmergasse 22
50676 Köln / Cologne
Germany

PRESENTED BY:
NORC at the
University of Chicago
Eric Weiss, PhD
Senior NORC M&E Analyst
55 East Monroe Street
30th Floor
Chicago, IL 60603
(312) 759-4000

NORC at the UNIVERSITY of CHICAGO

Table of Contents

Table of Figures

Table of Tables

Foreword: The Challenges of Impact Assessment

Roger Peltzer
Project Director, Competitive African Cotton Initiative (COMPACI)

The Competitive African Cotton Initiative (COMPACI) is funded by the Bill & Melinda Gates Foundation, the German Ministry of Cooperation (BMZ), the Aid by Trade Foundation and the Gatsby Foundation. The Initiative started in 2009 and has two primary objectives:

(1) To improve the livelihoods of 755,000 Smallholder cotton farmers in 12 African countries, primarily through the improvement of yields for cotton and also for the food crops planted by the farmers.

(2) To make the cotton production of these farmers more sustainable and to qualify it for verification under the Cotton made in Africa criteria (CmiA), thus giving African smallholder farmers access to the growing world market for sustainably grown cotton.

Donors, retailers engaged in selling Cotton Made in Africa, and also the general public are of course very interested to learn whether or not COMPACI/CmiA have a real impact on the livelihoods of the concerned farmers and on the "sustainability" of their cotton production and if so to what extent.

For this reason COMPACI/CmiA designed an ambitious M&E and verification concept from the start of their activities. We have learned, however, that the measurement of program impact is by far more complex then all implementing partners had thought at the beginning. The complexity starts with the scale of the project. Measuring the impact of a project that touches hundreds of thousand smallholder farmers in 12 different African countries is a challenge in itself. Additionally, the rather simple question of how much cotton a cotton farmer produces on his or her cotton field turned out to be far more difficult than one would think: In absence of GPS measurement, most farmers do not know their actual plot size and farmers' recall of the quantities of cotton produced and sold proved to be unreliable in many parts of Africa.

This is why the COMPACI-management together with its M&E contractor NORC conceptualized a new survey method to estimate yields accurately and collect representative information on critical indicators. The so called "representative sample based yield surveys" are supported by GPS devices and computer tablets and were first implemented in Zambia and Zimbabwe in 2013. Within these surveys, yields are estimated through counting cotton bolls on samples of the field and then scaling up to an estimate for the entire field instead of relying on farmer recall. Cotton field sizes are all measured by GPS. While there is still room for improvement on technical details, this survey methodology provides rather reliable yield results. It is the intention of COMPACI/CmiA to roll out these surveys to most of their partner countries in 2014 and thereafter conduct them annually. Thus a database will be created for most cotton producing African countries containing representative time series on cotton yields, applied agricultural techniques as well as some socioeconomic data.

The present report describes why and how the sample based surveys developed from the COMPACI M&E experience and explains their methodology. Additionally it provides an overview of the types of data collected during the two pilot surveys in Zambia and Zimbabwe. The first interesting results derived from these pilot surveys shows their value: They provide insight into project impacts, can spur discussion on effective agricultural training techniques among implementing partners, and they provide valuable information for the Cotton made in Africa verification.

Indication on project impact

A comparison of yields between the baseline conducted by NORC in 2010 (where plot sizes were already measured with GPS) and the Zambian "DZL" yields from the 2013 sample based survey in Central province show a yield increase of 23 % = 130 kg/ha. The data are not 100% comparable as the 2010 figures were based on farmer recall and the yield increase cannot accurately be attributed to COMPACI/CmiA because a control group is missing. Nonetheless, additional supplementary information supports the notion that it is valid to compare baseline and endline data and that COMPACI/CmiA had some influence on the yield improvement. First, the incidence of side selling, which has the most important effect on the reliability and accuracy of farmer recall data (or, more properly, the production levels reported by the farmers to the survey enumerators) was low in 2010; we therefore conclude that farmer "recall data" in 2010 were probably close to the truth. Second, climatic conditions in 2013 were unfavorable, thus could not cause rising yields. And third, farm gate prices for cotton were low in the 2012 season, hence farmers were probably not highly motivated to care for their 2013 cotton (which was part of the survey) in expectation of high prices.

Spur discussion on effective training methods among partners

The results of the survey show that the application of good agricultural techniques (e.g. regular weeding) or techniques to improve soil fertility (mulching, crop rotation, manure etc.) has positive effects on yields.

The picture is, however, not as clear when it comes to ripping, a method of soil preparation heavily promoted by COMPACI/CmiA. According to survey data, the acceptance of ripping in Zambia remains rather low, with somewhat higher figures in Central province (18% with DZL) and the data do not give a clear picture of whether or not ripping leads to higher yields. This finding should enable discussions among COMPACI implementing partners in Zambia and eventually lead to corrective action, if necessary.

On the other side, a rather positive feedback regarding the general COMPACI training for cotton farmers can be drawn from the Zambian survey. More than 75% of all interviewed farmers reported receiving training; this figure is higher in areas where COMPACI has been active for a longer period than in zones that only recently joined the program. This is a remarkable result, because almost 50% of Zambian smallholder cotton farmers go in and out of cotton from year to year and are thus most probably not so motivated to attend regular training.

Provide valuable information for Cotton made in Africa verification

The surveys provide some valuable information for the CmiA verification. They show, for example, that smallholder farmers in Zambia and Zimbabwe have cultivated areas of between 2 and 5 ha, of which 35-40 % are cultivated with cotton. Virtually all surveyed farmers practice crop rotation, mainly with maize and groundnuts, which is an important Cotton Made in Africa criterion.

It can also be seen that approximately 70% of the farmers receive training in scouting and threshold spraying, which is an important element of integrated pest management (IPM) and more than 75% of farmers practice 2 or more techniques for improving soil fertility - both important CmiA criteria.

One key claim of Cotton made in Africa is that improved livelihoods of cotton farmers lead to better primary school attendance of cotton farmers' children. Determining primary school attendance rates is therefore also part of the survey. Compared to the 2010 baseline, primary school attendance increased from 65% to more than 80% in Zambia, which is a very good result, even though it is not possible to attribute it to COMPACI/CmiA, as other factors probably also contributed to this positive development. In the future, sample based surveys will also determine the average duration of primary school attendance.

In total, it can be seen that the instrument of sample based surveys supplies valid information for the COMPACI management, for CmiA verification, and for the improved management of the cotton companies. The implementation of these surveys on a continuing annual basis with a majority of COMPACI/CmiA farmers is therefore expected to generate valuable data sets that serve a variety of purposes and stakeholders.

1.0 Introduction

The overarching objective of the COMPACI Program is to increase smallholder cotton farmers' income from cotton in the (currently) nineteen Sub-Grantee cotton companies in the ten African countries participating in COMPACI. Given that many smallholder farmers are constrained by the amounts of land or labor available to them, it is necessary to increase cotton production on the available land rather than focus on cultivating cotton on larger/more cotton fields. Income from cotton is to be increased primarily by providing training on cotton intensification along with other interventions to participating farmers in order to increase their cotton yield (kg/ha). Therefore, any assessment of the impacts and outcomes of the COMPACI Program must include some assessment of the changes in participating farmers' cotton yield over the life of the Program that are attributable to the COMPACI interventions. This will be done by comparison of yields estimated at later times to those calculated as part of the Baseline survey. A Control Group of cotton farmers will be used where possible, and other bases for comparison (perhaps the national statistics) used where Control Groups cannot be defined.

Note that the methods and evolutionary process (and the results) reported here apply primarily to COMPACI East and Southern African countries, although they also apply to Ghana and Benin. In both Burkina Faso and Cote d' Ivoire, the COMPACI Sub-Grantees have more structured national cotton sectors and cotton concession areas in which they have exclusive rights to buy cotton from farmers. They are also enabled to forward sell their cotton which consequently results in a greater need for accurate, advance yield estimates. In response to this situation, these Sub-Grantees have been using systems and methods similar to those recently piloted in Zambia and Zimbabwe and about to be introduced in Malawi, Mozambique, Ghana, and Benin.

Note that the discussion in this report is focused on *farmer-level yield*, that is, the cotton yields that are calculated for individual farmers. The reason for this deliberate decision by NORC to focus on farmer level yields is that this enables analyses of the relationships between different specific and collective cotton farming practices and yield. Also, if the survey from which the data were developed used a *statistically representative sample of farmers*, then these data can be aggregated up to produce a ginnery-level yield estimate. In contrast, as will be discussed later in this report, the ginneries are only interested in ginnery-level production and yield, which means that the yield estimates developed by ginneries and those derived from the NORC surveys cannot, in general, be compared, for reasons that will be explained later in this report (Section 2.3).

This report presents an overview and summary of the yield estimation problem and the various methods used over the course of the COMPACI Program to provide estimates of cotton yields at the farmer level. In addition to a discussion of these methods and their strengths and weaknesses, a more in-depth discussion is provided of the method currently being piloted and implemented, primarily in Eastern and Southern Africa (as noted, some of the West African COMPACI Sub-Grantees (cotton companies, or "ginneries") have already been using this method.) This newly implemented approach is needed in order to overcome problems noted in earlier methods and to ensure accurate, high confidence yield estimates that address the shortcomings of the earlier methods.

This remainder of this report is organized as follows:

- Section 2: The Evolution of Yield Estimates in COMPACI
- Section 3: Lessons Learned: The Problems in Cotton Yield Estimation in COMPACI
- Section 4: The Current Approach to Yield Estimation in COMPACI
- Section 5: Results from the Current Yield Estimation Method
- Section 6: The Way Forward
- Annex 1: Details of the In-Field Sampling Approach to Yield Estimation
- Annex 2: Questionnaire Used in a Zambia Pilot Survey

2.0 The Evolution of Yield Estimates in COMPACI

Because of the criticality and centrality of accurate, high confidence yield estimates to COMPACI, yield estimates have been made from the very beginning of COMPACI Monitoring and Evaluation (M&E) activities. As problems with these estimates or the methods used have become apparent, lessons learned were applied to the next, improved, iteration of these surveys.

This section presents a brief overview of the evolution of yield estimates in this series of surveys.

2.1 Baseline Surveys

Starting in June 2010, the first COMPACI Baseline survey was conducted in Zambia. These surveys all used *representative samples;* that is, the statistical results from these surveys could be projected onto the COMPACI sub-grantee farmer populations as a whole. Control Groups were also used to provide the basis for later attribution of changes in yield to COMPACI. Also, following the COMPACI protocol of using local institutions to collect survey and other data, NORC employed local research institutions in each COMPACI country to interview the farmers and to do the field area measurements. NORC developed the questionnaires and provided training to the local research institution teams on all aspects of these surveys.

The *production* part of the yield estimates was based on farmer recall of production for the previous season; in general, these surveys were conducted as soon after the harvest as possible in order to ensure more accurate farmer recall. (Unfortunately, the first of these Baseline Surveys, in Zambia, was carried out before the harvest and buying season were completed, resulting in a significant percentage (25%) of the surveyed farmers not having sold their cotton by the time they were interviewed.)

The farmers' cotton plot sizes were in all cases measured with hand-held GPS units to ensure accurate plot sizes. Note that, because of crop rotation, many farmers alternate which piece of land is planted in cotton. Farmers may have two plots, one with cotton and one with maize, with the crops on each plot switched back and forth to avoid soil exhaustion. The Baseline Survey protocol was to ask about production and to measure the area of the plot used in the *previous* season.

The necessity for using GPSs to measure cotton field areas was evident from the beginning; in the first Baseline Survey in Zambia, the surveyed farmers were also asked how big they thought their cotton fields were. When these results were compared to the actual field sizes as measured with the GPS units, only 8% of the surveyed farmers knew the correct area of their cotton fields to within $\pm$ 10%. In general, farmers tended to overestimate (or, at least, to overstate) their field sizes by a considerable margin, a pattern seen in virtually every COMPACI survey that included both self-reported (by the farmers) and measured field sizes.

Upon review of the yield data that were derived from these Baseline Surveys and comparison of these data to the company-level yield estimates being reported by the ginneries, it was evident that the results were generally not comparable, although at that time the reasons for this significant divergence in estimates were not known. (See Section 3.3 for a more complete discussion of this point.) What was evident, however, was that an improved approach to yield estimation was needed during the COMPACI project and for the Endline survey; these intermediate yield estimates and those from the Endline survey will be compared to the Baseline estimates and will thus be used, together with Control Groups or other comparable yield data for non-COMPACI farmers to assess COMPACI impacts on cotton yields.

2.2 The 2011 Malawi Yield Survey

At the start of 2011 Malawi was believed to be one of the COMPACI countries where the yield estimates reported by the Sub-Grantee (GLCC) were most questionable and likely to be least accurate or reliable. Therefore, Malawi was chosen as the venue for the next iteration in yield estimation surveys. In this case, the yield estimates were based on the farmers' self-reported amounts of cotton produced (kg). Because the Baseline Surveys had shown that using external (non-ginnery) enumerators was problematic because these outside enumerators were generally unfamiliar with cotton farming, making them less effective in interviewing cotton farmers for these surveys. For the 2011 Malawi Yield Survey, Sub-Grantee (GLCC) extension staff were used as enumerators. This study and the associated were conducted in late 2011-early 2012.

A secondary objective of this survey was to determine the impacts on of input selection on cotton yield; specifically, the impacts of using good *vs* poor cotton seed and also the use *vs* non-use of other chemicals for the cotton. This was a much smaller survey involving a total of 163 GLCC (the Malawi Sub-Grantee) farmers. This was *not* a statistically representative sample due to the relatively small number of farmers. The complete results of this survey were documented in an official COMPACI report, *Malawi Comparative Yield Study: Analysis Results,* May 2012.

This survey was also the first one in COMPACI in which farmers' reports of cotton production were compared to ginnery records of cotton purchases from these same farmers.

For almost two- thirds of the farmers (62%) the measured cotton plot sizes were found to be smaller than the registered sizes, and that for 20% of the farmers, the sizes were the same (in some cases this is because no plot size was registered, so GLCC retroactively accepted the *measured* size as the *registered* size.) GLCC suggested that many of the registered sizes were set by the farmers at higher levels than needed in order to qualify for more inputs (some of which may have been diverted to their maize). As noted, GPS units were used to measure the field sizes; on one half-hectare cotton plot the Garmin GPS Map60 unit measured the area within about 6.5%, compared to the Garmin GPS Map60csx unit that measured this area within about 1.5%. (From 2012 onward only Garmin GPS Map62csx or eTrex units have been used; these are at least as accurate as the Garmin GPS Map60csx.)

Note in this survey 18 (11%) of the surveyed farmers (11%) showed *no* cotton sold to GLCC *but reported non-zero amounts of cotton produced in 2011*. The most likely interpretation of this is that these 18 farmers (along with a number of the other surveyed farmers) side-sold their cotton (i.e., sold their cotton to someone/some company other than the company to whom they were contracted) and thus sold none to GLCC. The most likely reason for this was to avoid repaying the credit advanced to them by GLCC.

In order to not lose 11% of what is already a small sample, for this analysis cotton yield was calculated as follows:

YIELD = KG purchased by GLCC / Measured area (ha) *if kg purchased by GLCC > 0.0*

KG reported by the farmer / Measured area (ha) *if kg purchased by GLCC = 0.0*

The problems of side selling and “reverse side-selling” (bulk selling) immediately became apparent in the survey results and in attempting to do an analysis of the achieved yields. One farmer reported producing 615 kg of cotton in 2011, but according to GLCC records, he sold *no* cotton to them in that year. His 2011 cotton plot was measured to be .546 hectares. His yield (kg/ha) using the amount of cotton *sold* to GLCC was 0.0 kg/ha, but if his *reported production* is used instead, his cotton yield in 2011 was 1126 kg/ha.

A phenomenon that is related to side-selling but is in a sense the “reverse” situation occurs when a farmer, not wanting to repay the credit advanced by a ginnery, gives his cotton to another cotton farmer to sell to that ginnery on his (the original farmer’s) behalf. This situation is indicated in the data by the amount of cotton sold by one farmer being so much that the calculated yield is highly doubtful or even impossible. For one of the farmers in the 2011 Malawi survey; both he and GLCC agree that he produced and sold to GLCC 1126 kg of cotton. But, his measured cotton plot was only .37 ha, leading to a calculated yield of 3033 kg/ha, a highly doubtful number. His *registered* plot size was not much bigger at 0.4 ha, so even if this larger plot size were used to calculate yield, the results would still be the not-believable 2815 kg/ha. It is very likely that this farmer was selling not only his cotton but that of another farmer as well.

To further explore the relationship between the *purchased* and the *reported* cotton production of the farmers, the *ratio* of the *amount of cotton (kg) purchased by GLCC-to-the amount of cotton reported by the farmer* was calculated. So, for example, 11% of the farmers had a ratio of exactly 0.0; i.e., they sold *no* cotton to GLCC in 2011. Further, 35% of the surveyed farmers, sold *no more than* 0.50 of the cotton they reported producing to GLCC.

Perhaps more significantly, 60% of the farmers sold less cotton to GLCC (according to GLCC records) than the farmers produced, according to what they told the survey enumerators (who were GLCC field staff). This statistic suggests a very large amount of side-selling. Also, the data show that for only 39% of the farmers, the amount they reported producing agrees with the amount that GLCC records show they purchased.

Note that two different values for the sizes of the farmers'2011 cotton plots were available; the sizes as measured with GPS devices, and the sizes that were registered with GLCC. As GLCC noted, many of the sizes registered with them were set somewhat arbitrarily, or were deliberately exaggerated to enable the farmer to get additional inputs, raising further doubt about their accuracy and reliability. The data show that for almost two- thirds of the farmers (62%) the measured sizes are smaller than the registered sizes, and that for 20% of the farmers, the sizes are the same (in some cases this is because no plot size at all was registered, so GLCC accepted the *measured* size as the *registered* size.) As stated above, GLCC suggested that many of the registered sizes were set by the farmers higher than needed in order to qualify for more inputs (some of which may have been diverted to their maize).

Among the conclusions drawn from this survey was the recognition of the likely large magnitude of side-selling and bulk-selling, both of which confound any attempt to accurately measure yields at the farmer level. These problems were compounded by using ginnery extension staff as the enumerators, since this increased the likelihood that farmers would not truthfully report their cotton production to the company with whom they had a contract.

In view of the above experiences, COMPACI Management directed that an improved yield estimation survey and study be conducted in Malawi in the following year (2012).

2.3 The 2012 Malawi Yield Survey

Based on the lessons learned in the 2011 Malawi Yield Survey some modifications were made to the questionnaire and survey protocols used for the 2012 survey. The most significant of these were the following:

- External (non-GLCC) enumerators were used to interview the farmers
- Questions intended to specifically capture data on side-selling and bulk-selling were added to the questionnaire. In particular, the 2012 survey included questions for the farmers regarding *total amount of cotton harvested, the amount of cotton sold to GLCC, the amount of cotton sold to other parties (side-selling), the amount of cotton farmers sold for other farmers ("bulk-selling.")*
- There was a recognized need to explicitly include questions in the survey to consider other factors that are likely to influence cotton yield. Such factor include *planting date vis a vis the date of the first rains, the extent to which the farmers practiced the cotton farming methods promoted by GLCC under COMPACI, and other cotton farming practices used by the surveyed farmers.* Consequently, The 2012 survey instrument included questions regarding numerous other factors that may affect yield at the individual farmer level. These included:
    - Land preparation time (month, part of month) (to be compared to the date of first rains in that area)
    - Planting time (month, part of month)
    - Use of chemicals and amounts used, if any, and frequency of use
    - Station gap spacing
    - Number of seeds per station
    - Thinning and the number of plants left per station
    - Gap-filling and replanting
    - Good Agricultural Practices (GAP)
    - Harvest time (start month/period, primary month/period, and end month/period)
- Detailed farmer-level purchase data were provided by GLCC and merged with survey data
- Rainfall data for all surveyed areas to be obtained from the Malawi Met Service and provided by GLCC to allow comparison of the date of the first rains to the onset of planting dates in each area.
- Different input usage regimes were *not* used in determining the survey sample of farmers

- This survey considered the yields obtained by different *types* of farmers. Specifically, the 2012 study used a sample of 300 farmers that were grouped into one of five "farmer categories." The farmer categories, and their numbers and proportion in the overall farmer sample are defined in the following bulleted definition:
  - *Lead Farmer:* A farmer used to demonstrate the promoted cotton farming techniques and to train other farmers (28, 9%)
  - *Best Farmer:* A farmer that achieves higher than average cotton yields but is not a Lead Farmer (30, 10%)
  - *Follower Farmer:* A farmer that has been trained by a GLCC Lead Farmer under COMPACI (118, 39%)
  - *Ordinary Farmer:* A Farmer that has not been trained by a GLCC Lead Farmer under COMPACI and may not have had any prior relationship with GLCC (121, 40%)
  - *Seed Farmer:* A farmer that grows cotton in part to sell the seeds commercially (3, 1%)
- Farmer subgroups did *not* include an atypical percentage of better (Lead) farmers (as happened in the previous survey) in order to calculate more realistic average yields across farmer subgroups and across GLCC as a whole

Some of the principal results from this survey are summarized below. The complete results of this survey are presented in a COMPACI report, *Malawi Comparative Yield Study: Analysis Results from the 2012 Survey*, 18 December 2012.

In this survey, when the farmers were asked how *much cotton they harvested in the 2011-2012 season*, the average amount reported was 757 kg with a median amount of 500 kg. In contrast to these data, when GLCC was asked how much cotton they had purchased from the surveyed farmers, the average amount reported by the farmers was 578 kg with a median amount of 392 kg. These two average production amounts show the large difference between the amounts of cotton reportedly produced by the farmers and the amounts purchased from them by GLCC, and is indicative of large amounts of side-selling and "bulk selling" (one farmer selling cotton for several farmers and thus showing the sale of more cotton than he/she harvested).

Planting date was found to have a significant effect on yield; the average yield for farmers that planted their cotton before the last third of November is somewhat erratic, whereas, the maximum average yield was obtained for farmers that planted their cotton in the last third (end) of November; the figure shows a steady decline in average yield for farmers that planted their cotton after that period.

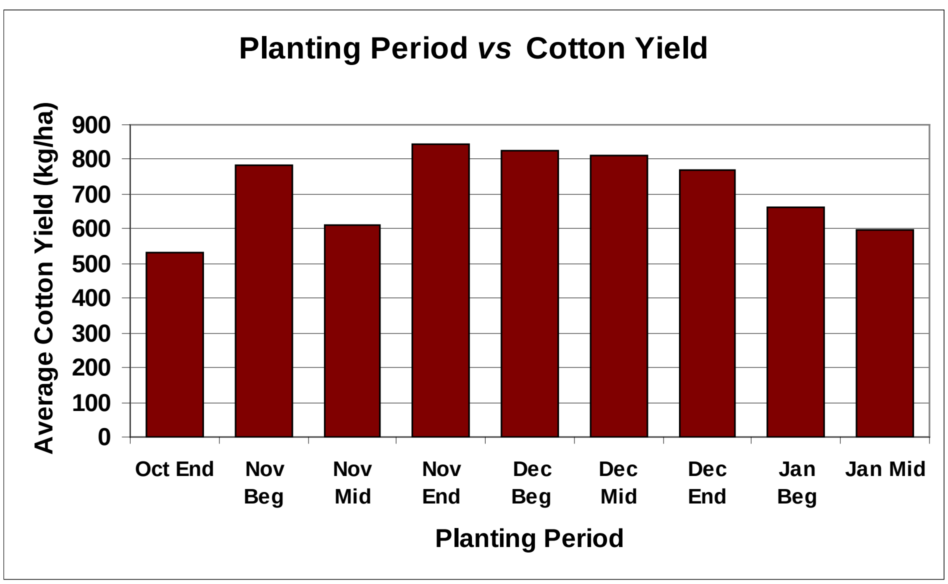
What explains the erratic behavior in average yield *before* late November? There are two factors:

(1) In the Shire Valley, where much of this survey was implemented, many farmers practice *"dry planting;"* that is, they plant their cotton *before* the rains. If the rains are delayed, this can reduce the germination rate, require replanting later than is optimal, and hence the yield of the cotton.

(2) In 2011, the rains began in late October and then stopped, only to start again in late November, from which date they continued normally. Therefore, planting before end November was, in some sense, equivalent to the *dry planting* mentioned above, with the same results.

The pattern described in the above statements can be seen graphically in Figure 2-1 which explores the relationship between *planting date* and *average yield.* The decrease in cotton yield with increased delay in planting after the rains begin is clearly seen in this figure. This has implications for the current Field-based sampling method used in the current approach to yield estimation and will be discussed in Section 5 of this report.

Figure 2-1: The relationship between planting date and cotton yield



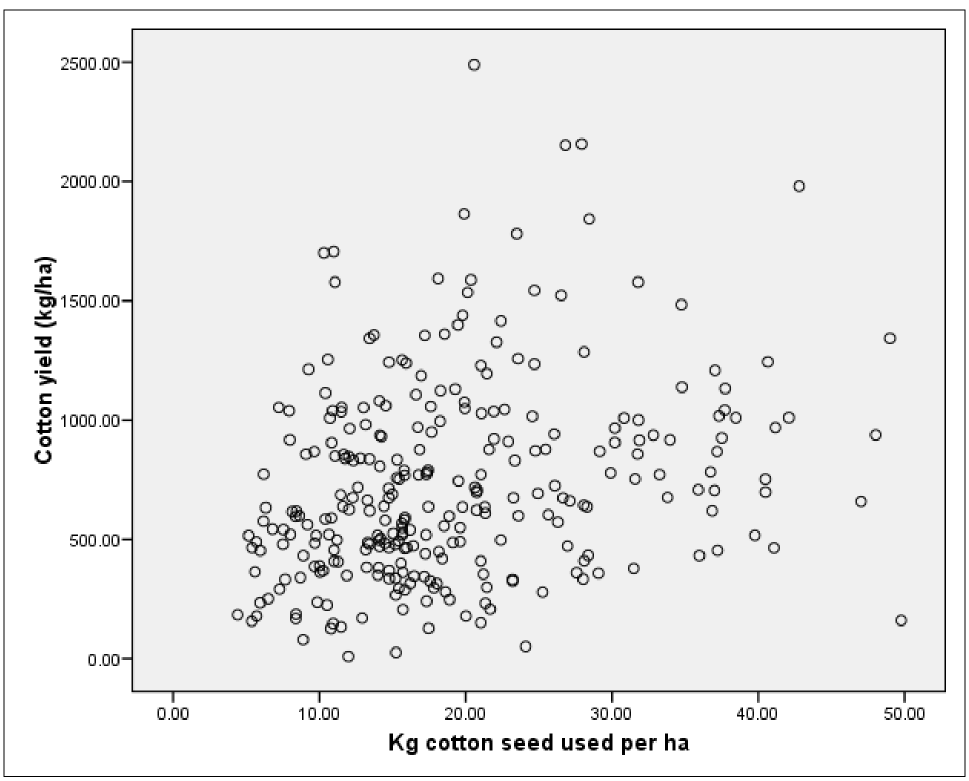
Another key result of this survey confirmed that an important factor affecting yield is the amount (kg) of seed used per hectare. GLCC promotes the use of 10-12 kg of cotton seed per hectare of land to be planted in cotton. However, there are several factors that can affect the amount of seed actually used by the farmers; among these factors are:

- The *row separation* used by the farmer (distance between ridges on which the cotton is planted)

- The *station gap* used by the farmer (distance between planting stations along one ridge). Although GLCC recommends a gap of 45 cm, the surveyed farmers had gaps varying between 30 cm to 80 cm with an average of 49 cm
- Seed germination rate and the need to fill in any gaps where no seeds germinated.
- The number of seeds used per planting station. The surveyed farmers used between 1 to 10 seeds, with an average of 4.3 seeds per station.
- The need to replant cotton, for whatever reason. Approximately 90% of the surveyed farmers replanted their cotton due to the early start and subsequent temporary failure of the rains.

Because of these variable factors, the relationship between *amount of seed used* and *yield* may be fairly complex and may exceed the recommended amounts by a wide margin. The survey data presented in Figure 2-2 show that, up to about 30 kg of seed per hectare, yield increases with the amount of seeds used. After about 30 kg/ha, yield declines with increased amounts of seeds. Part of this may be explained by the fact that most of the farmers had to replant their cotton thereby using about twice as much seed as would otherwise have been necessary. Also, it is possible that some of these farmers do not follow some of the promoted practices for cotton farming. The sample size used for this survey (300 farmers) does not permit detailed statistical analysis with disaggregated subsets of farmers.

Figure 2-2: Cotton yield and *amount of seeds per hectare*



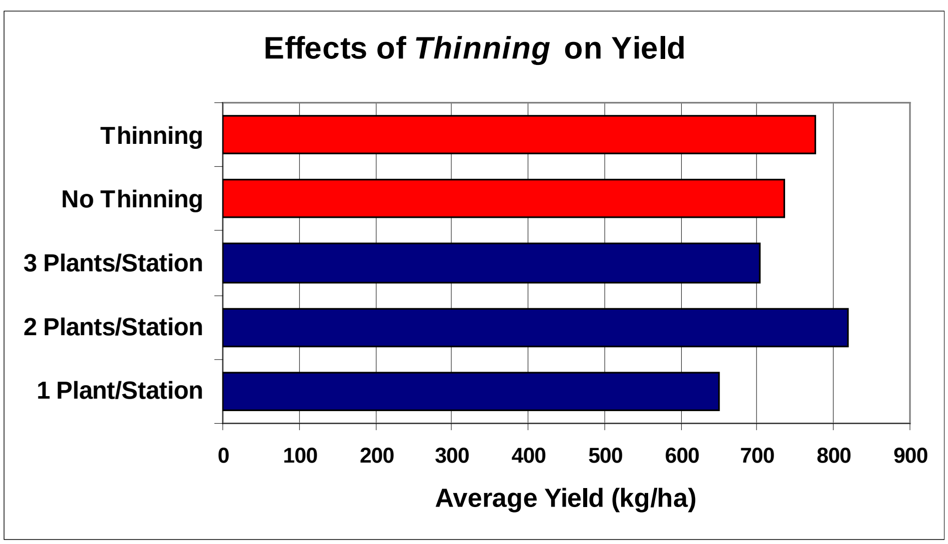
Another factor that was identified during this survey as having a substantial effect on cotton yield was the practice of *thinning*. *Thinning* is the process of removing excess cotton plants from stations with more than one growing cotton plant. This is a practice promoted by GLCC (and other cotton companies) under COMPACI as a way to increase yield as more cotton is produced overall when the plants are not in such extreme competition for moisture and nutrients. However, to some farmers the removal of perfectly good cotton plants to increase yield is counterintuitive, and many resist following this practice.

Of the surveyed farmers in this 2012 Malawi study, 270 (90%) reported doing thinning of their cotton plants. For these farmers, *thinning* increased their yield, on average, about 5%, from about 735 kg/ha to 777 kg/ha over farmers that did not *thin*. Another issue related to *thinning* is *how many plants are left per station.* That is, depending on how many seeds are initially placed in each station, the number of viable plants growing in each station will vary. GLCC promotes leaving two (2) plants per station after *thinning.*

Figure 2-3 shows the average yield from *thinning* or *not thinning*, and also the average yield for different numbers of plants left after *thinning*. This figure confirms that, at least for the surveyed farmers, the optimal number of plants to leave after *thinning* is two.

Other less critical factors determined to be affecting yield and other results from this survey may be seen in the complete survey report, cited above.

Figure 2-3: Cotton yield and *thinning*



3.0 The Problems in Yield Estimation in COMPACI

The previous section discussed the evolution of yield surveys in the COMPACI. From this series of surveys, some key lessons were learned and conclusions drawn regarding the methods used to estimate yield. These lessons learned and conclusions, discussed in this section, coupled with the approach used by some of the West African Sub-Grantees led to the current approach to yield estimation that is discussed in Section 4.

In order to accurately estimate the yield of any field crop, both the field size (the area, measured in hectares, acres, square meters, etc) and the amount of production (kg, metric tons, etc) must both be known with reasonable accuracy. In estimating cotton yields in much of Africa, accurate measurements or estimates of either or both of these parameters have proven to be problematic as explained below.

3.1 Cotton Field Area

In general, smallholder farmers do not know the sizes of their fields with any accuracy because they have no means to accurately measure the fields. In some sense, the field size as an abstract number is not inherently important to them; only the amount of production from the field and the income earned from this production are of any importance to the smallholder (although as this author's experience can confirm, when the fields are measured with a GPS, many or most of the farmers are eager to know the actual field size).

The Eastern and Southern African ginneries to which these farmers are contracted typically estimate the area of their farmers' cotton fields by either using a tape measure to measure the perimeter or by walking around the perimeter ("pacing the field") and then trying to use simple geometry to estimate the area from the perimeter length and shape. This method is only, at best, reasonably accurate for rectangular fields. For the irregularly shaped fields found in many African countries (and specifically used by many COMPACI smallholders), such methods are completely incapable of producing even reasonably accurate estimates of the field area.

In an experiment carried out by the author during the 2011 Yield Estimation Survey done in Malawi, a test was made of the accuracy of an improved variant of this basic pacing approach. In this modified approach, the field staff were trained to use tape measures to measure the length of the field "sides" and compasses to measure the bearing direction of each side (this enabled calculation of the angle at each field "corner.") Where the field's edges were curved, the staff had to decide where to consider the end of each “side” so as to approximate the actual curved side by a series of shorter straight sides and then measure the length of these straight field sides.

In this experiment, six two-person teams consisting of ginnery field staff measured the length of the sides of the same 0.5 ha irregularly shaped field, and used a compass to determine the bearing directions of the field sides. These data were recorded. An online application was used by NORC to calculate the area resulting from the measurements of each of the six teams. Figure 3-1 shows the different shapes of the field that resulted from the six teams’ measurements. Note the significant variation in the apparent shapes and sizes that resulted from these different sets of measurements.

Figure 3-1: Different shapes of the same cotton field as assessed by six separate teams

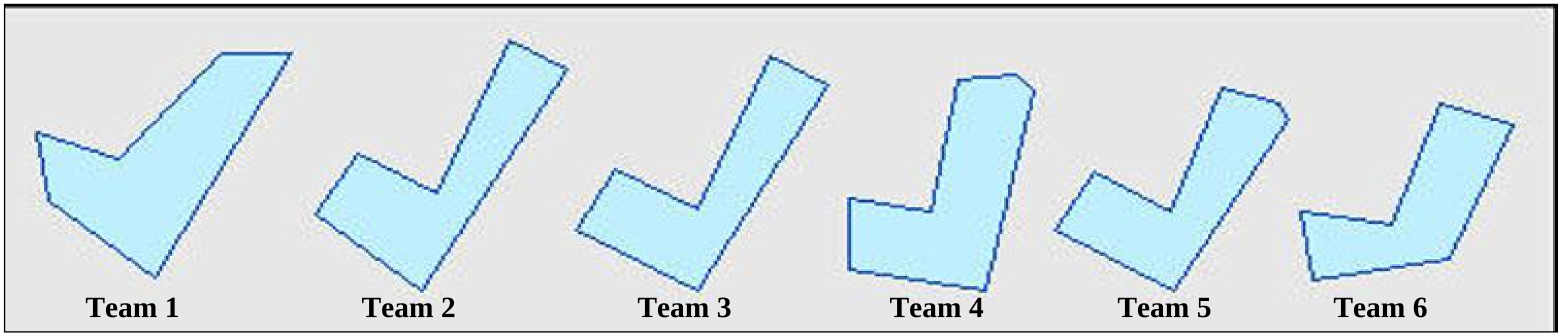


Table 3-1 presents the data underlying these different shapes. Note that the different teams did not even agree on how many sides to use in measuring the perimeter, and that the lengths and bearing direction angles of analogous sides also varied widely. Finally, note that the calculated area of these different estimates of the same field varied from 0.345-0.511 ha, a variation of 0.166 ha, a variation of 48%. A more complete discussion of this experiment is presented in the COMPACI report *"Malawi Comparative Yield Study: Analysis Results,"* 25 May, 2012, also by this author.

Table 3-1: Team measurements of common cotton plot

Enum. Team	S1 Bearing (Degrees)	S1 Length (Meters)	S2 Bearing (Degrees)	S2 Length (Meters)	S3 Bearing (Degrees)	S3 Length (Meters)	S4 Bearing (Degrees)	S4 Length (Meters)	S5 Bearing (Degrees)	S5 Length (Meters)	S6 Bearing (Degrees)	S6 Length (Meters)	S7 Bearing (Degrees)	S7 Length (Meters)	Area (ha)
1	310	60.0	350	34.2	110	38.3	40	69.5	90	31.1	210	128.0			0.451
2	310	60.0	30	35.0	120	41.0	22	83.0	120	29.0	220	124.0			0.388
3	300	60.3	20	35.5	120	37.7	30	65.6	80	24.1	120	12.5	220	126.9	0.406
4	280	60.4	360	35.0	100	36.6	10	67.7	83	25.5	140	11.7	210	124.2	0.430
5	300	60.0	30	33.6	120	38.5	20	65.7	106	25.7	150	9.1	210	125.4	0.511
6	260	60.3	350	34.8	100	40.0	20	65.0	110	32.7	220	125.6			0.345

"S1" refers to the first side, "S2" to the second side, etc. Note that all teams started at the same point along the field edge, so the side numbers are comparable

Figure 3-1 and Table 3-1 jointly show that measuring the cotton plots with tapes and compasses, a more accurate method than simply pacing the field, is still not a good way to produce accurate, consistent estimates of the field area. What is not seen in these data is the difficult and tedious work required to use this method to get the basic measurements. And, it should be noted that calculating the areas of these irregular shapes is also not an easy job, even if computer-aided calculations are used. Therefore, ginnery records of estimated field sizes based on some form or variation of pacing the farmers' fields are not considered sufficiently accurate for COMPACI impact assessment purposes.

The other method used by some East and Southern African ginneries to estimate field sizes uses the amount of inputs provided; specifically, the amount (kg) of cotton seed given to the farmers on credit by the ginneries. Since this amount is, for some cotton companies, determined by the amount of land the farmer has declared that (s)he will plant in cotton, estimating the field size by the amount of seed given is circular reasoning since the amount of seed is itself often based on the farmer's estimate (or claimed) field size.

Gracious Hamatala, the former Dunavant Limited Zambia (DZL) COMPACI liaison, claimed in a discussion with the author that overall, as a total estimate of the land planted in cotton (ha) by *all of their farmers*, 0.85 x (Number of 1 ha seed packs distributed) had historically proven to be a reasonable estimate. However, as Mr Hamatala also conceded, this estimation method says next to nothing about the size of any individual farmer's field. Since the seed is only given in packets meant for 1 ha, any farmer whose claimed, registered, or estimated field size is less than about 1.5 ha will get the same 1 ha seed pack. Consequently, for purposes of estimating yield at the farmer (vs ginnery) level, or for use in COMPACI, this method is also unacceptably inaccurate.

The conclusion from all of the forgoing is that the only reasonably accurate method of measuring farmers' field sizes is with the use of hand-held GPS units. This has been the NORC practice from the start of the COMPACI Programs, and was, as already noted, first used by NORC in all of the Baseline Surveys. Originally, the Garmin GPS Map60 was used, but as these are no longer in production, more recent NORC surveys have use the newer and more accurate Garmin GPS Map62 or the equally accurate and easier to use Garmin eTrex 20 or eTrex 30 units.

3.2 Cotton Field Production

Getting accurate estimates of the amounts (kg) of cotton produced by individual farmers is also problematic, although for different reasons. There are some alternative approaches to determining farmer-level production.

The most direct assessments of production can be made by simply asking farmers what their total cotton production was; since their cotton is weighed by the buyer at the time of sale, and the farmer is usually given a receipt, the exact quantity of cotton is known. The first problem with this approach is that the farmer may deliver cotton on several occasions throughout the season, and so may not remember the total amount that (s)he sold.

For example, of the 817 Alliance Zambia farmers surveyed in the 2013 Yield Estimation Survey, slightly less than half (392 – 48%) only made one delivery of cotton to Alliance, 126 farmers (15%) made two deliveries, 39 (5%) made three, and 27 (3%) made between four and eleven deliveries. (Also according to Alliance records, 228 (28%) of the surveyed contracted farmers sold *no* cotton to Alliance, indicating either a high level of side-selling and/or a very poor season in which farmers may have abandoned their cotton fields.)

If in fact some of these farmers side-sold their cotton, then even if they remember the true amount produced, they are unlikely to report that to the cotton company to whom they are contracted (since cotton company extension staff conducted the 2013 Yield Estimation Surveys.)

Thus, if the farmers either truly do not remember the total production amount correctly or they side-sold some cotton, the amounts of production that they report will be inaccurate, possibly deliberately so. Thus, farmer recall is an unreliable method of gaining an accurate estimate of farmer level production.

A second method of estimating farmer level production is by extracting this information from ginnery records. However, the ginneries only have records of the amounts of cotton *purchased* from individual farmers, which may not be the same as the amounts *produced* by these farmers. If the setting is a country without concession areas or with inefficiently controlled concession areas where side-selling and bulk-selling may be common, then ginnery records may not be accurate at the individual farmer level in those cases either.

To see this, consider Figure 3-2, which represents data from the 2012 Yield Estimation Survey in Malawi. In this 2012 Malawi survey outside enumerators were used in the hope that farmers would be more truthful about this than if they were being interviewed by GLCC extension staff). The *ratio of the amount of cotton bought by GLCC to the amount of cotton reported produced by the farmers* was calculated and the percentage of surveyed farmers in each ratio band was calculated. Any value of the ratio less than 1.0 represents a case in which GLCC (the Malawi COMPACI sub-grantee) bought less cotton than the farmer reported producing, whereas any case for which the ratio is greater than 1.0 represents a farmer who sold more cotton to GLCC than he reported producing, a possible case of bulk selling. Finally, any farmer whose calculated ratio was exactly 1.0 sold exactly as much cotton to GLCC as (s)he reported producing.

Figure 3-2: Cotton production vs sales by farmers, Malawi, 2012

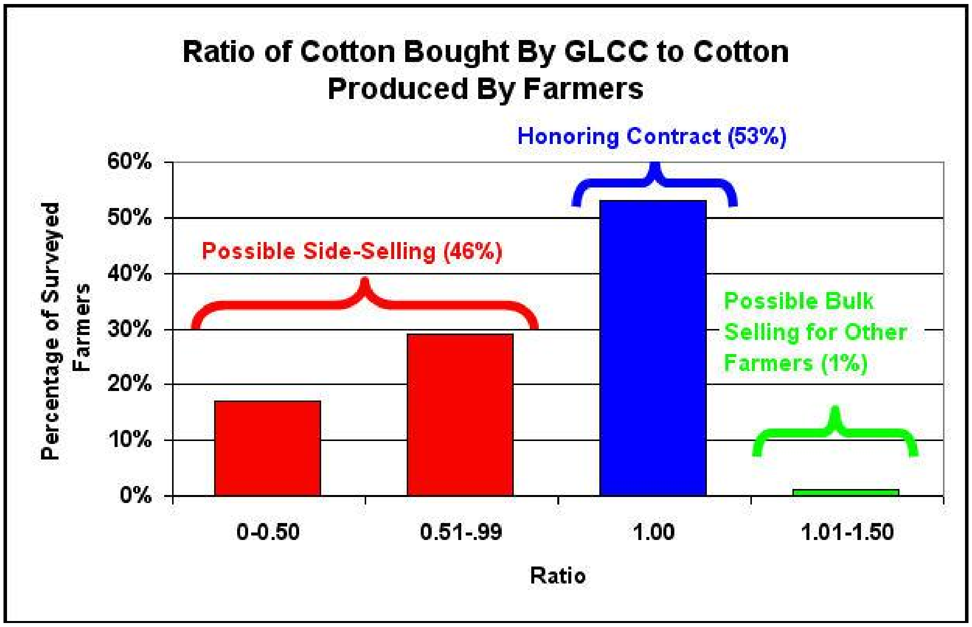


Figure 3-2 shows the percentage of surveyed farmers in different ratio bands. This figure shows that approximately 46% of the 300 farmers interviewed in this survey may have side-sold some or all of their cotton (ratio <1.0), 53% of the farmers appear to have honored their contract, and 1% may have bulk-sold cotton for other farmers in addition to themselves.

This result proves the general principle that, in areas in which side-selling and bulk-selling are likely or confirmed to occur, ginnery records will not provide high confidence, accurate estimates of farmer-level production.

3.3 Comparison of COMPACI and Ginnery Yield Estimation Methods

It might be expected that for those NORC surveys that use a statistically representative sample of farmers, that the farmer-level yields could be suitably weighted and aggregated to produce a ginnery-level estimate of yield and that this ginnery level estimate would be comparable to the yield estimates produced by the ginneries themselves.

Indeed, this was attempted in with data from the Baseline surveys, but it was found that there was still no agreement between the survey-generated ginnery –level yield estimates and the estimates produced by the ginneries. As more experience was gained with both the surveys and with the data produced by the ginneries, the reasons for this disparity in the estimates has become clear. The primary reasons for the difference in estimates are summarized below:

(1) Ginneries report *total cotton purchased / total estimated hectarage*. The immediate problems with this approach in terms of comparison to farmer-level results are:

  (a) This approach does not account in the calculations for side selling or bulk selling, so in countries where these occur, purchased amounts often do not equal produced amounts;

  (b) The ginneries for the most part have no farmer-level accurate estimates of plot sizes but are using either some form of pacing of plots and/or inputs provided to get their estimates (some West African Sub-Grantees such as Faso Coton and Seco Coton are exceptions to this in that they systematically use GPSs to measure farmer plot sizes).

(2) Ginneries report a *company* (i.e., *ginnery*) level figure whereas NORC calculates the average farmer level yield; these two calculations typically will not agree (see Table 3-2, below).

(3) NORC estimates are based on *production*, NOT on *purchased* amounts of cotton

(4) The ginneries, by virtue of their estimates of total hectarage under cotton typically include the estimated hectarage of farmers that abandoned their fields because of poor rains and very low production as happened in the 2012-2013 season (Dunavant had this problem and reported including some 28,000 farmers that delivered no cotton) which of course lowers their reported figure considerably

The basic reason that the overall ginnery-level yield will not agree with the average of farmer-level yields as might intuitively be expected can be surmised from the simple example presented in Table 3-2 which compare the results for a sample of just three farmers. The core reason for the difference in results is that the ginnery approach implicitly weights farmers' yields by the size of their cotton fields whereas the NORC approach takes the simple, unweighted average of farmers' yields.

Table 3-2: Comparison of yield calculation methods

Farmer	Production (kg)	Cotton Plot Size (ha)	Yield (kg/ha)
1	1,000	10	100
2	2,000	5	400
3	3,000	3	1,000
Average Yield (NORC Method): (100 + 400 + 1,000) / 3 = 1,500 / 3			**500**
Average Yield (Ginnery Method): (1,000 + 2,000 + 3,000) / (10 + 5 + 3)= 6,000 / 18			**333.3**

Note, however, that if suitable weights are developed for a statistically representative sample of farmers whose individual yields have been calculated, then it should be possible to produce a NORC estimate of total *production* for all farmers contracted to that ginnery. However, this estimate will still not compare to the ginnery-generated yield estimates because, 1); they are based on *purchases* (in lieu of production) for any COMPACI Sub-Grantee that does not have an operative concession area, and, 2); except for the West African Sub-Grantees mentioned earlier, the total estimated hectarage will not be as accurate as the total hectarage scaled up from the GPS measurements of the sample farmers' fields.

A better idea of the divergence in the different yield estimates, taken from the report of the 2012 Malawi survey is presented in Table 3-3.

Table 3-3: Estimates of GLCC farmer cotton yield (kg/ha)

Cotton Yield (kg/ha)	Data Source	Applicable Season	Comments
324	GLCC Report to DEG	2009/2010	Non-representative samples, questions about accuracy of plot size estimates
566	NORC Baseline	2009/2010	Representative sample. Yield calculated *using plot sizes reported by farmer* – these should be comparable to GLCC plot size data for those farmers
844	NORC Baseline	2009/2010	Representative sample. Yield calculated *using plot sizes measured by survey enumerators using GPS devices*
465	GLCC Report to DEG	2010/2011	Non-representative samples, questions about accuracy of plot size estimates
850	GLCC Report to DEG	2011/2012	Non-representative samples, questions about accuracy of plot size estimates
1081	NORC 2011 Yield Survey	2010/2011	Non – representative sample purposefully selected by GLCC to assess impact of inputs on yield. The issue of side-selling was not dealt with in the farmer interviews, with the overall result of uncertain credibility and likely bias of yield results
775	NORC 2012 Yield Survey	2011/2012	Non-representative sample selected by GLCC to include specific types of farmers in approximately correct proportions to overall population. Side-selling addressed

The above data show the confusion that characterizes attempts to accurately and credibly measure cotton yield in Malawi. The principal factors which bias and distort the data are the prevalence of side-selling and bulk selling, both within GLCC and across companies, the known inaccuracies in plot size estimates at the individual farmer and company levels, and the non-representative nature of the samples of farmers from which the yield estimates are calculated.

Finally, note that these problems are not unique or confined to Malawi: Table 3-4 shows the different yield estimates made for Zambia in the 2012-2013 season; the different assumptions behind the results, the methods used, and the sources for these estimates are also listed here. All of these data were from the 2012-2013 season; almost all of them show significant improvement over the 2009-2010 season Baseline yield of 538 kg/ha; however, it should be noted that this yield was based on farmers' self-reported cotton plot sizes rather than GPS-measured sizes. Also, the Baseline value was computed for only those farmers that grew cotton in that season and not all farmers in the survey grew cotton that season (7% did not grow cotton that season).

Table 3-4: Different yield estimates for Zambia, 2012-2013 season

Organization	Data Source	Affected Area	Estimated Yield (kg/ha)	Notes
DZL	DZL Reports to COMPACI	All Operating Provinces	450	▪ All operating areas of Zambia ▪ Company-level statistic ▪ Farmer field sizes estimated ▪ *EARLY SEASON ESTIMATE*
NORC-DZL	NORC Report	All Operating Provinces	288	▪ Estimated purchases of 33,000 MT ▪ Includes 28,000 farmers with 0 kg deliveries ▪ *POST-HARVEST ESTIMATE*
NORC-DZL	Yield Estimation Survey	Eastern Province Only	871	▪ Eastern Province only ▪ Representative sample of farmers ▪ Farmer-level statistic ▪ Fields measured with GPS ▪ DZL-provided estimates of average boll weights (3.9 g)
NORC-DZL	Yield Estimation Survey	Central, Lusaka Provinces	759	▪ Central, Lusaka Provinces only ▪ Representative sample of farmers ▪ Farmer-level statistic ▪ Fields measured with GPS ▪ DZL-provided estimates of average boll weights (3.9 g)
NORC-DZL	Yield Estimation Survey - *MODIFIED*	Central, Lusaka Provinces	661	▪ *Use average Alliance ZA boll weight = 3.4g* ▪ *Estimated yield 5% bigger than Alliance yield*
NORC-Cargill	Cargill Report to COMPACI	Eastern Province only	800	▪ Projection based on Sept 2012 data
NORC-Alliance	Yield Estimation Survey	Central Zambia Provinces	630	▪ Central, Lusaka Provinces only ▪ Representative sample of farmers ▪ Farmer-level statistic ▪ Fields measured with GPS Alliance-provided estimates of average boll weights (3.0 – 3.8 g)
Cotton Farmers Association of Zambia, (CAZ)	NORC Meeting with Cotton Development Trust (CDT)	National	321	▪ Estimated area cultivated: 325,642 hectares ▪ Estimated Production: 104,433 metric tons
CDT	NORC Meeting with Cotton Development Trust (CDT)	Magoye (Southern Province)	N/A	▪ Average Boll weights by seed cotton variety: 1) Chureza: 6.5 grams per boll 2) CDT-5: 5.6 grams per boll 3) CDT-2: 5.4 grams per boll ▪) Basal (Max-crop) and top dressing (CAN-LAN) applied

4.0 The Current Approach to Yield Estimation in COMPACI

One overriding conclusion from all of the preceding surveys is that yield estimates should not be based on farmer recall in any form or on cotton company purchase records if these options can be avoided. The current COMPACI approach to yield estimation discussed in this section is the direct response to that finding.

The approach currently being developed and recently piloted in Zambia and Zimbabwe is in many respects similar to the yield estimation methodology used by Faso Coton in Burkina Faso and by Seco Coton in Cote d' Ivoire. These COMPACI Sub-Grantees have concession zones and also forward-sell their cotton so they have both the means and the motivation to accurately estimate their farmers' total yield *before the cotton is harvested.*

This methodology in question uses hand-held GPS devices to measure the area of the cotton fields and is coupled with in-field sampling to estimate the production of that field and farmer interviews to understand the cotton farming practices used y them. This work takes place *before* the harvest but when the plants are mature enough that the good bolls on a sample of the plants can be recognized and counted.

An overall summary of the methodology is presented below; see Annex 1 for a more detailed description of this methodology, including the equations used in it.

(1) The Sub-Grantee sends their current farmer list to NORCS
(2) A representative sample of the Sub-Grantee's contracted farmers is determined (by NORC)
(3) Sub-Grantee extension workers are used as survey enumerators and are trained by NORC to do the interviews and the sample fieldwork.
(4) For each of the sample farmers, the following activities take place:
    (a) The area of the cotton field is measured with a hand-held GPS unit
    (b) Depending on the measured area of the field between 5-8 randomly selected 3 meter strips of cotton rows are sampled (Note: for 2014 5 meter sample strips will be used.)
    (c) For each 3 meter sample strip, the distances to the row on the left and on the right are measured.
    (d) The number of plants in each 3 meter strip is counted.
    (e) The total number of viable cotton bolls (bolls that will produce cotton) in each 3 meter sample strip are counted.
(5) Using the GPS-measured area of the field and the data collected from the 3 meter sample strips, the overall numbers of cotton plants and bolls are estimated.

(6) The field production is estimated by multiplying the estimated number of bolls by the average boll weight provided by the Sub-Grantee for that area/planting date/variety.

(7) The yield for the field is estimated by dividing the estimated production (kg) by the measured area of the field (ha).

(8) As part of the overall survey, the farmer is interviewed regarding his/her cotton farming experience, cotton farming practices, and other issues of interest to COMPACI and/or CmiA (e.g., primary school attendance, other crops grown, etc). See Annex 2 for a copy of the questionnaire that was used in one of the pilot surveys in Zambia. (Note: the actual surveys were done using computer tablets, but the training was first done using paper questionnaires and then repeated using the tablets.)

The principal reasons for asking about cotton farming practices as part of this yield survey are: 1), To understand which practices the farmers are using and to compare these to the training provided by the Sub-Grantee to see which parts of the training are being followed and which parts ignored, and; 2), To be able to relate yield to specific cotton farming practices and *how* those practices are being implemented. For example, it is useful to know not just that the farmer *practices thinning* and how that affects yield, but to also see how *the number of plants left per station after thinning* affects yield (to illustrate this point in particular, see Figure 2.3)

To date (December 2013), this survey methodology has been piloted in Zambia (four surveys) and in Zimbabwe (two surveys). Current COMPACI plans call for this survey to be more widely implemented in 2014; specifically it is expected to be implemented again in Zambia and Zimbabwe and for the first time in Tanzania (four Sub-Grantees), Malawi (one Sub-Grantee), Mozambique (One Sub-Grantee), Benin (one Sub-Grantee), and in Ghana (one Sub-Grantee). Additionally, planning calls for both Faso Coton and Seco Coton to add the interview portion of this survey to their existing yield survey methodology in order to capture the data indicated above and to ensure consistency with the other COMPACI Sub-Grantees. When/if Cameroon and/or Ethiopia join COMPACI, it is expected that the new Sub-Grantees will also implement this survey.

5.0 Results from the Current Yield Estimation Method

This section presents some of the results from the four pilot surveys carried out in 2013. Section 5.1 presents results regarding *yield*, while Section 5.2 presents results relating to specific cotton farming practices and their relationship to yield and to showing the value, to the ginneries, of these surveys. One example of the value of these surveys to the ginneries is in demonstrating the value of using GPS devices to measure farmers' field sizes as a means of better determining the proper amounts of inputs (specifically cotton seed) to be given to the farmers.

Note that in all of the results presented in this section, the estimated yields were *only calculated for farmers that in fact grew and harvested cotton.* When the enumerators were visiting the selected sample farmers to interview them and to do the in-field sampling of their cotton fields, any farmer that did not have a cotton field with growing cotton was omitted from the survey and a "replacement farmer" was used in their place. This, in and of itself, means that the estimated yields produced by these surveys will be higher than ginnery estimates which included the hectarage of farmers that sold no cotton to them.

5.1 Cotton Yield

The yield estimation methodology presented in Section 4 showed that, in essence, the estimate of a field's production is based on first estimating the total number of cotton bolls in the field and then multiplying this number by the average boll weight for that area/planting date/variety. This observation makes clear the absolutely critical dependence of the accuracy of the yield estimate on having a correct boll weight; an error of X% in the assumed boll weight translates directly into an error of X% in the estimates of the field's production and yield. This issue is discussed more fully in Section 6.1, but for now it suffices to note that the accuracy of all estimates of production and yield are critically dependent on the accuracy of the boll weights provided by the ginneries.

Figures 5-1 and 5-2 show the distribution of estimated yields for Cargill (Zimbabwe) and Alliance (Zimbabwe), respectively. The average yields estimated for these ginneries' farmers were 662 kg/ha (Cargill) and 686 kg/ha (Alliance).

Figure 5-1: Distribution of estimated cotton yields for Cargill Zimbabwe

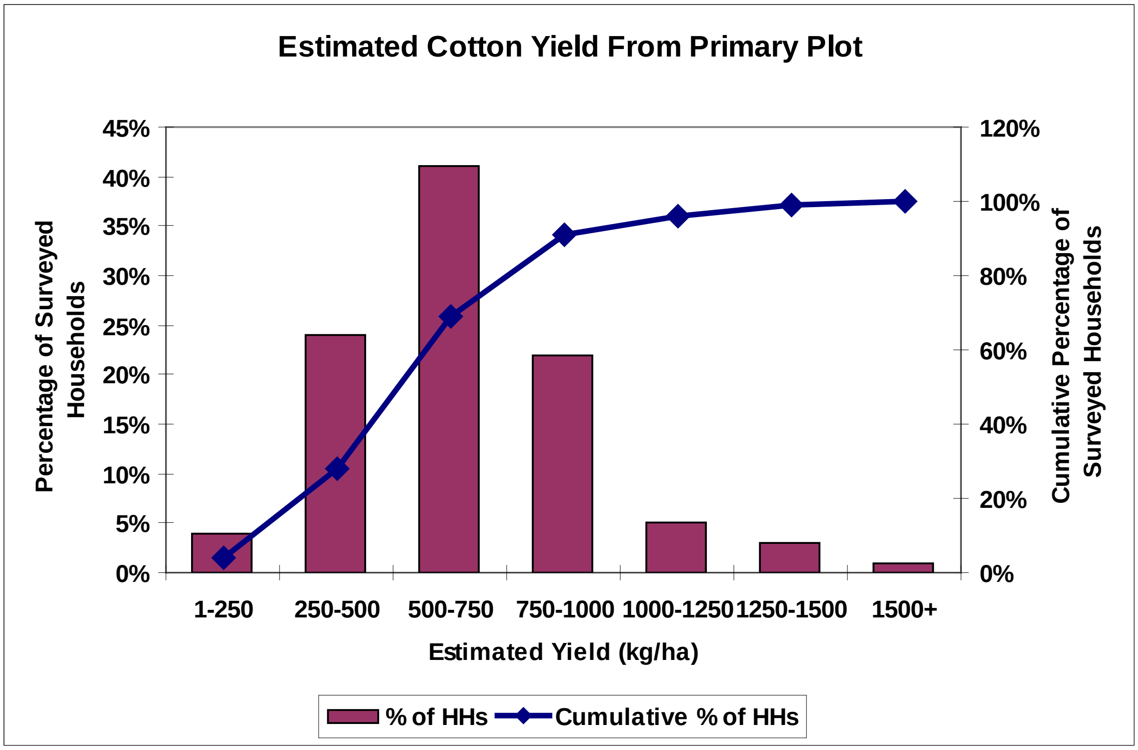
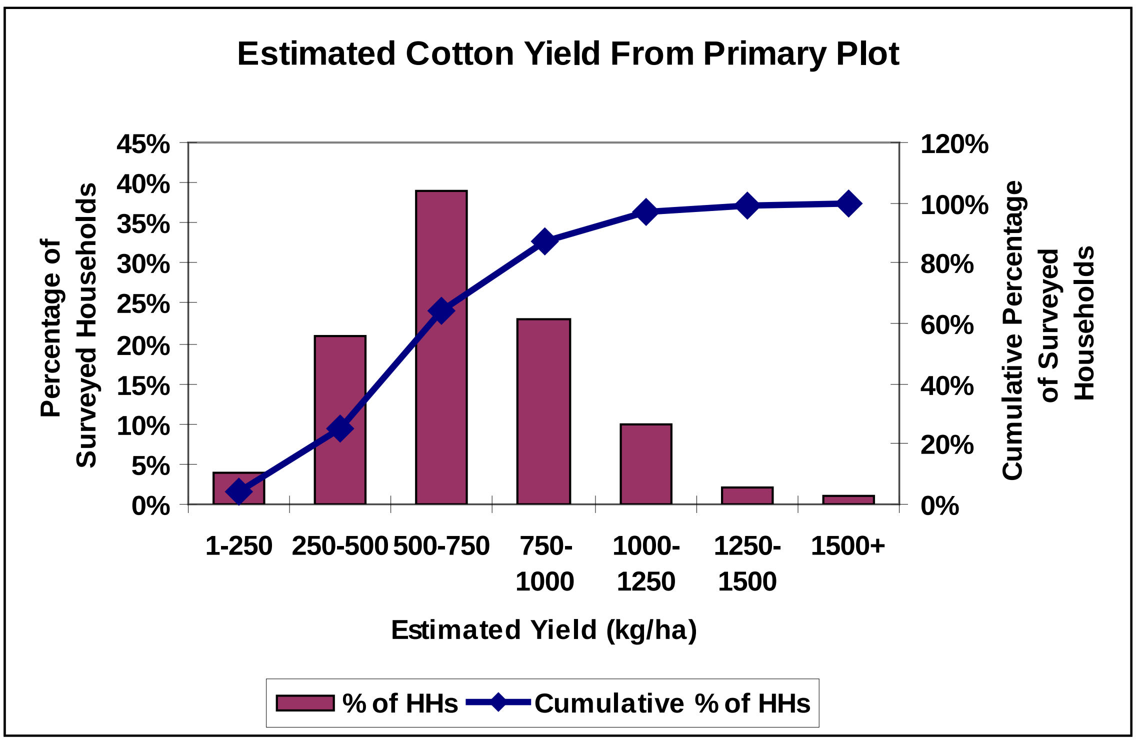


Figure 5-2: Distribution of estimated cotton yields for Alliance Zimbabwe



These two ginneries have largely overlapping operating areas and so the similarity in their average estimated yields and the distributions of those yields is not surprising. However, both ginneries developed the final boll weights that were used in these calculations by weighing samples of harvested bolls, so it *is* surprising that the boll weights used by these two ginneries were so different. Cargill boll weights varied between 2.5g-6.4g with an overall average of 4.2g while Alliance boll weights only varied between 2.0g-3.0g. The similarity in estimated yields coupled with the disparity in boll weights suggests that some other factors were at work and that Cargill plants may have had fewer, larger bolls compared to Alliance plants.

In Zambia, the Dunavant Zambia Limited (DZL) and Alliance have overlapping operating areas in Central Zambia, so it is useful to compare their average estimated yields at the farmer level. Figures 5-3 and 5-4 present these results for DZL (Central) Zambia and Alliance Zambia, respectively. The average estimated yields were 684 kg/ha (DZL) and 538 kg/ha (Alliance).

Figure 5-3: Distribution of estimated cotton yields for DZL (Central) Zambia

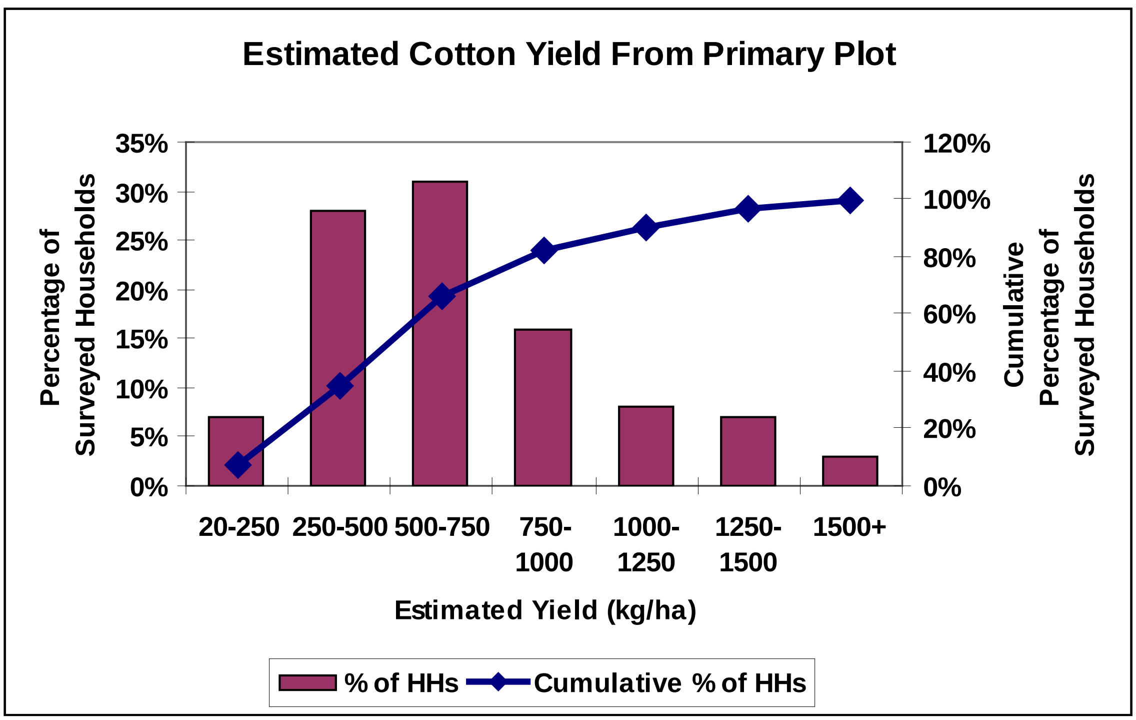
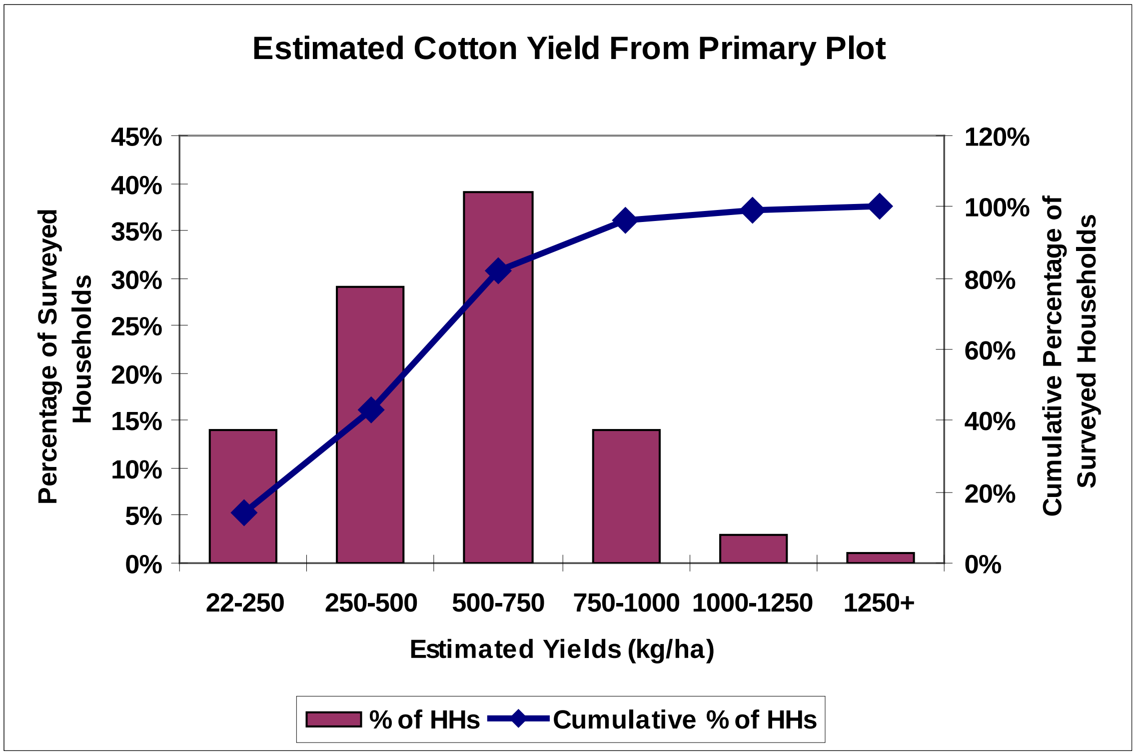


Figure 5-4: Distribution of estimated cotton yields for Alliance Zambia



In this case the sharp difference in yields may be explained by the boll weights used; DZL suggested that 3.9g should be used for all operating areas (in Central as well as Eastern Zambia) while Alliance provided boll weights that varied between 3.0g – 3.8g with an overall average of 3.4g . If this average Alliance boll weight of 3.4g is used to estimate the yield for the DZL farmers instead of the 3.9g boll weight provided by DZL, then the average DZL yield drops to 661 kg/ha, only 5% higher than the estimated average Alliance yield.

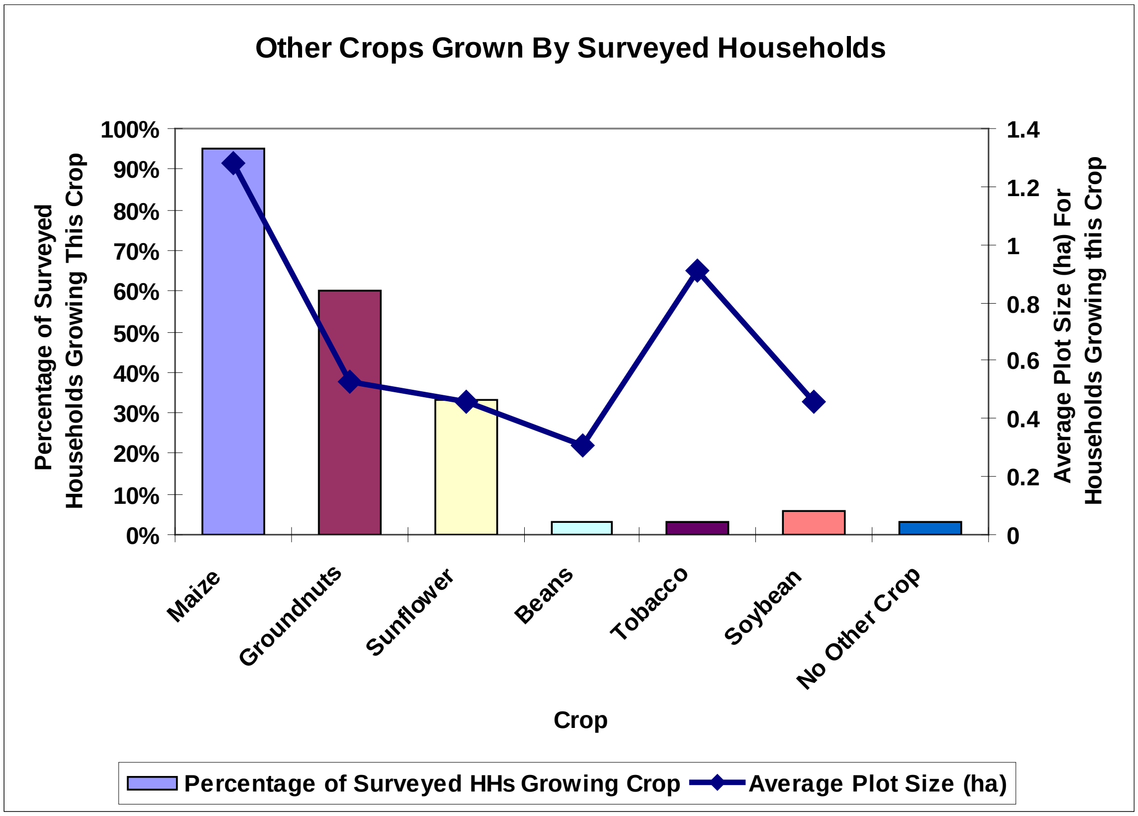
This last result suggests that the DZL boll weights were too high in at least some of the Central Zambia areas. This result also points out how critically dependent on the assumed boll weight the estimated yields are in this methodology.

5.2 Cotton Farming Practices and Yield and Other Results

This section presents select findings on the relationships between the estimated yields and different cotton farming practices as well as other results of interest.

One such other topic of interest is which other crops are grown by the surveyed households. Figure 5-5 presents the results from DZL (Eastern) Zambia. In this figure both the percentage of surveyed farmers growing a crop and the average plot size for that crop, *calculated for only those households that grow that crop*, are presented. This figure shows that, by far, *maize* is the most commonly grown other crop (95% of surveyed households grow maize) and that the average (self-reported but unmeasured) size of the maize plot is about 1.3 ha. In contrast to this, *groundnuts* are grown by about 60% of the households but the average plot size is 0.5ha and *tobacco* is grown by only 3% of the surveyed households but the average plot size for this crop is about 0.9 ha.

Figure 5-5: Other crops grown by DZL (Eastern) Zambia households and average plot sizes



Some of these other crops are used by some of the farmers growing them as rotation crops for cotton. Figure 5-6 shows the percentages of surveyed farmers from Alliance Zimbabwe that use some of these crops as rotation crops with cotton. Not surprisingly, *maize* is the most common rotation crop for these farmers.

Figure 5-6: Rotation crops for cotton for surveyed Alliance Zimbabwe households

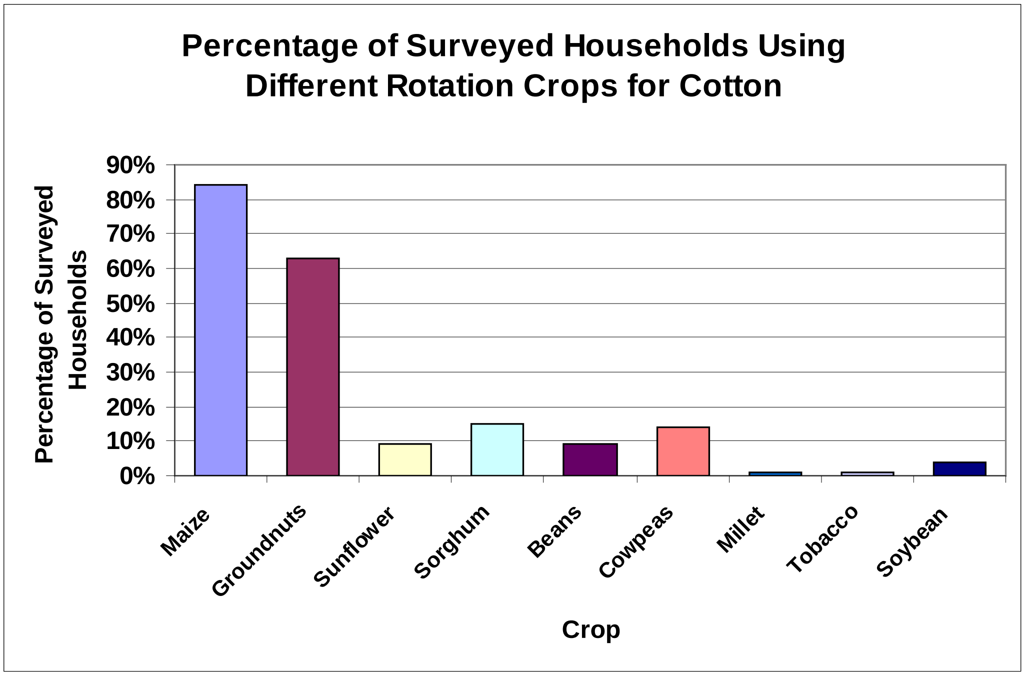


Figure 5-7 shows the percentages of surveyed DZL (Central) Zambia that use different select soil fertility methods.

Figure 5-7: Soil fertility methods used by surveyed DZL (Central) Zambia households

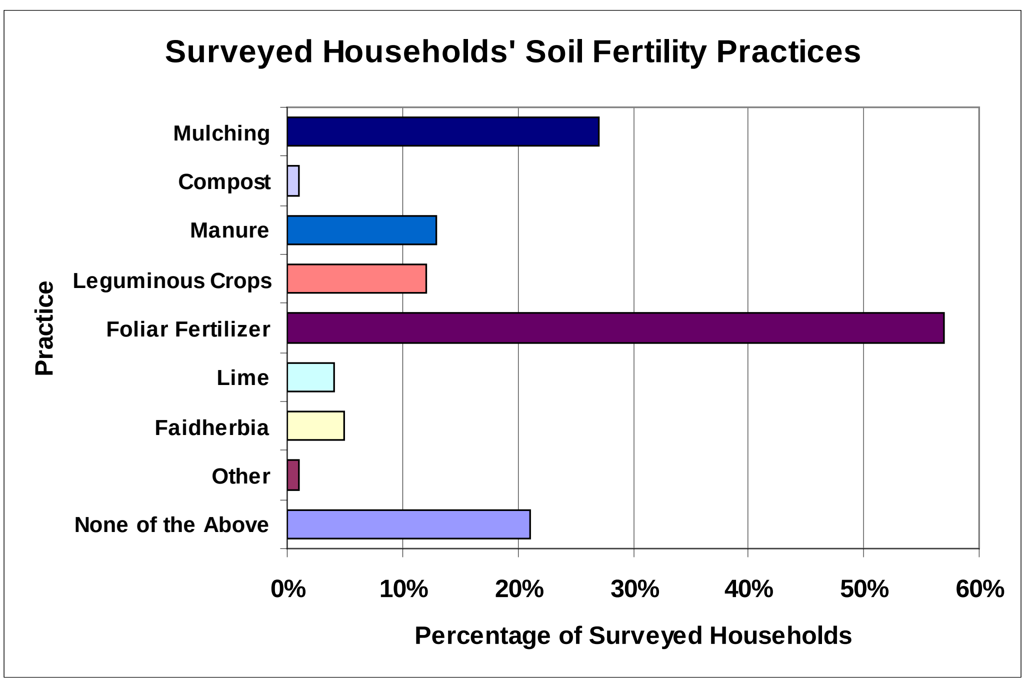


Figure 5-8 shows the number of different soil fertility methods, including crop rotation, that these same DZL (Central) Zambia households reported using. Finally in this series, Figure 5-9 shows the relationship between average cotton yield and the *number* of soil fertility methods used. This figure includes a (red) trend line and shows the clear benefit to yield of using multiple soil fertility methods (the most commonly used of these was *crop rotation)* which 80% of the surveyed households reported using on *all* of their cultivated land and another 11% reported using on *some* of their cultivated land.

Figure 5-8: Number of soil fertility methods used by DZL (Central) Zambia households

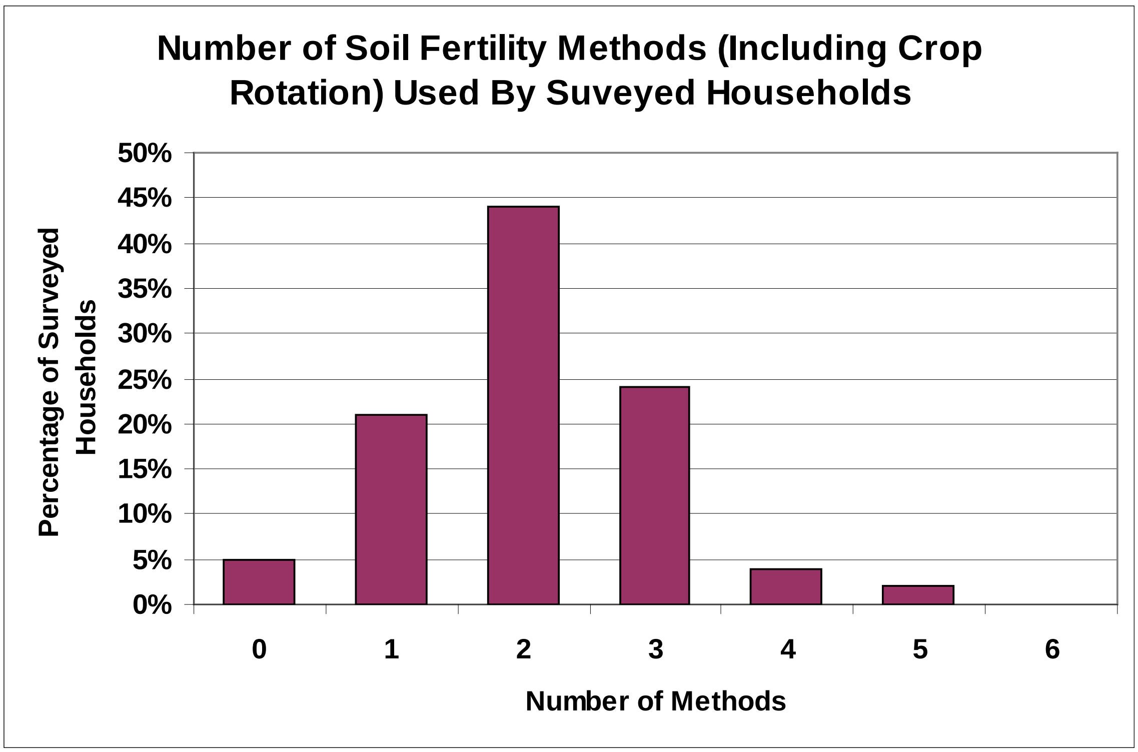
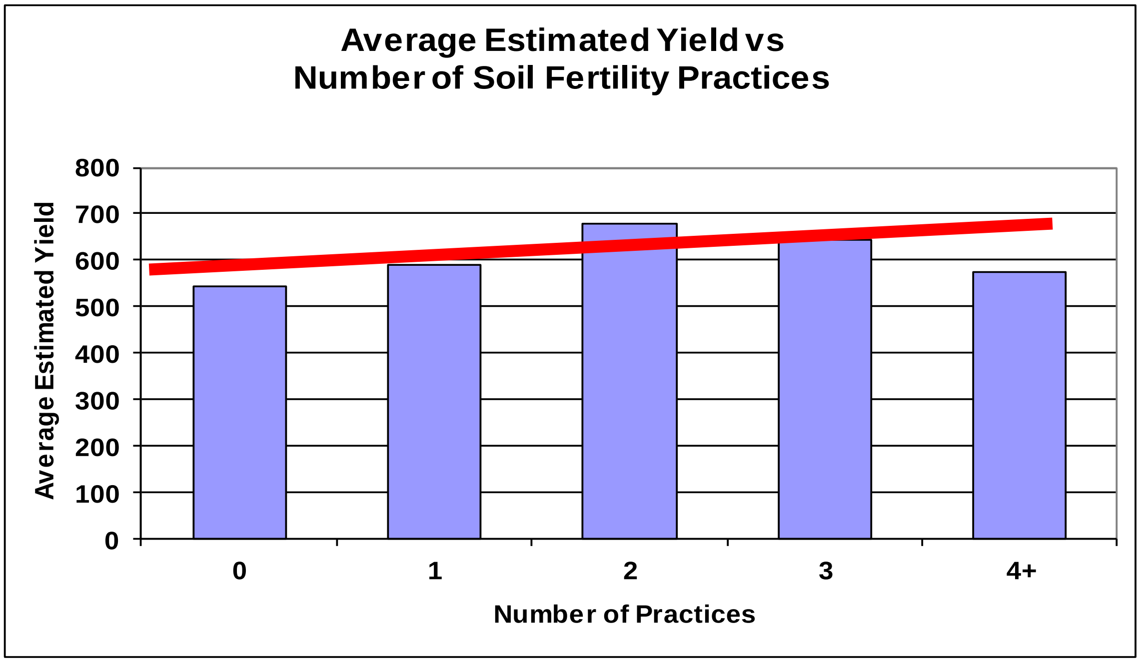


Figure 5-9: *Cotton yield* and the *number of soil fertility practices* used, DZL (Central) Zambia



Of the surveyed Cargill Zimbabwe farmers, 95% responded to the questions about *scouting*; of these farmers, 94% reported scouting between one and twenty times during the 2012-2013 season. The distribution of the number of times scouting was done is shown in Figure 5-10.

Figure 5-10: The number of times *scouting* was done in 2012-2013 season, Cargill Zimbabwe

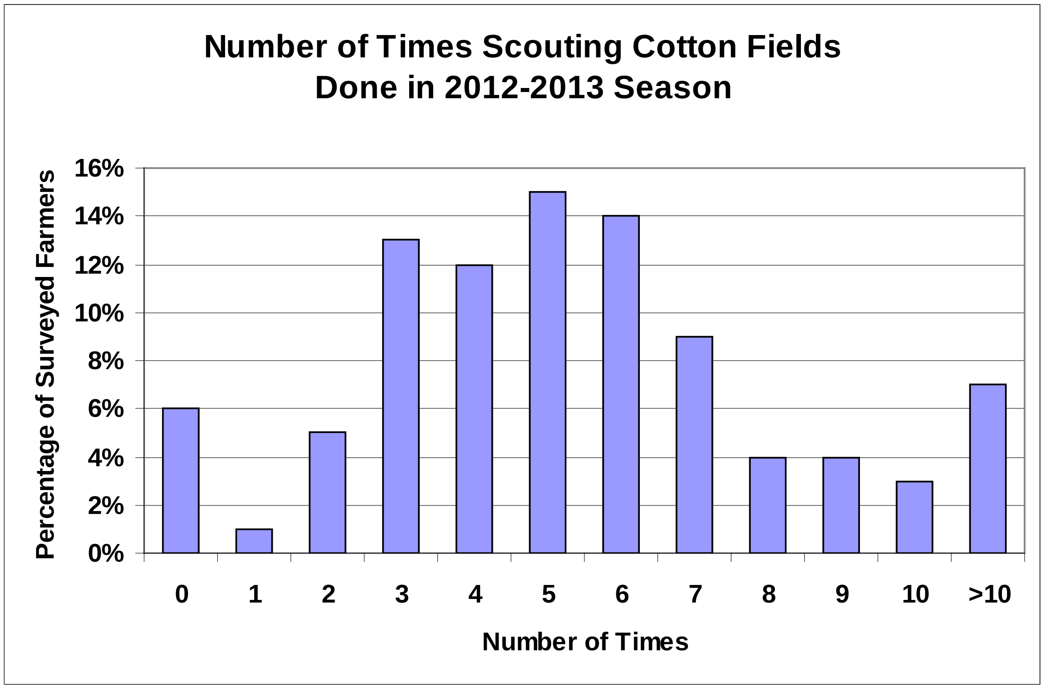
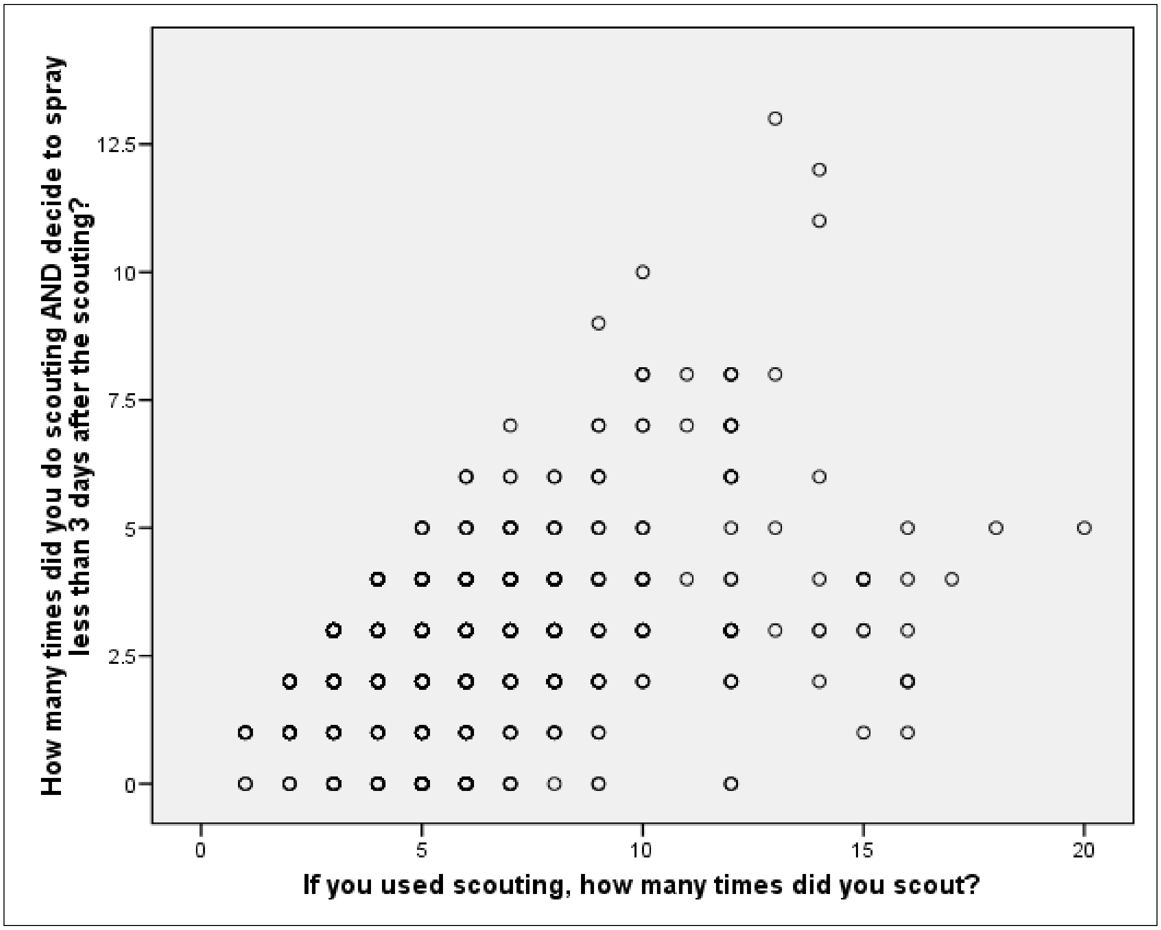


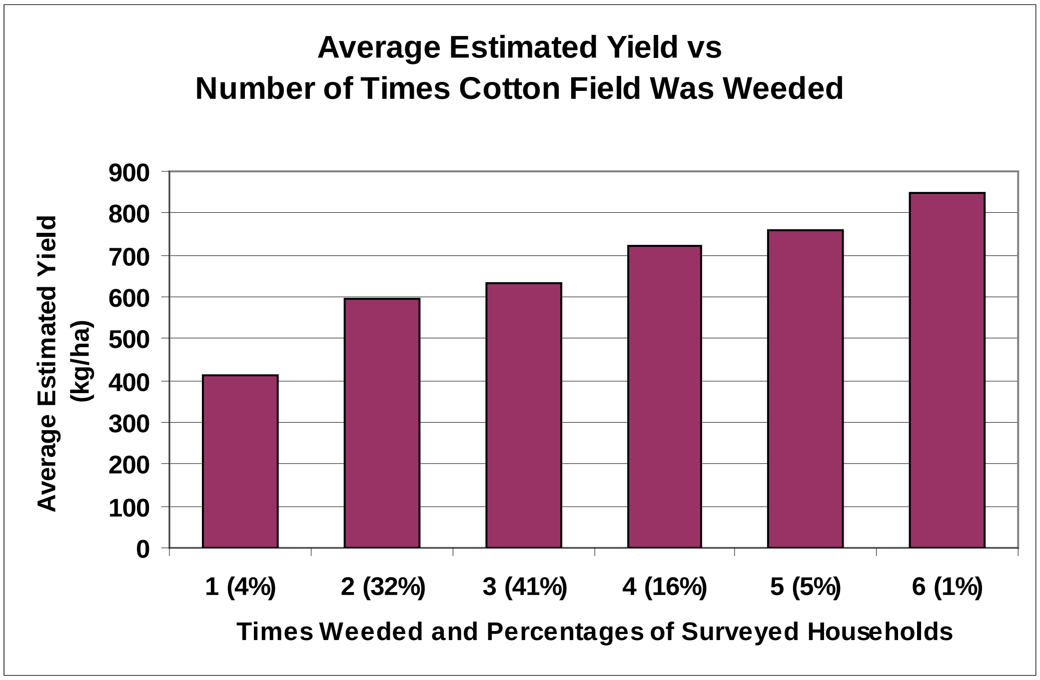
Figure 5-11 shows, for the same Cargill Zimbabwe farmers, for each number of times scouting was done, how often farmers had to spray afterwards based on the scouting results. This figure shows that the number of times scouting was required varied considerably from farmer to farmer and may reflect farmer practices, the care with which scouting was done, or the diligence of the farmers in spraying when required.

Figure 5-11: Scouting and the need for spraying afterward, Cargill Zimbabwe



The impacts on yield of more frequent weeding are apparent in Figure 5-12, taken from the survey results for DZL (Central) Zambia. Note that this figure shows, in the labels for the horizontal axis, not only the *number of times* the cotton field was weeded and the average yield achieved by surveyed farmers that weeded that number of times, but it also shows the *percentage of surveyed households that weeded that many times.* So, for example, 32% of the surveyed households weeded twice during the season and had, on average, a yield of 600 kg/ha whereas only 16% of these households weeded four times but had an average yield of 721 kg/ha. The single household that reported weeding six times had an even higher yield, but since this group included just one household, not too much should be made of this statistic.

Figure 5-12: Cotton yield and the number of times weeding was done in 2012-2013, DZL (Central) Zambia



The questionnaire used for the interview part of this survey asked about the usage of pesticide chemicals used against pests that attack cotton (*Bollworms, Aphids,* and *Stainers*). Specifically, the questionnaire asked about the number of times these chemicals were used throughout the 2012-2013 season by the surveyed households, and numbers and sizes of liquid and solid containers for these chemicals that were used by them. These data were combined with the measured areas of the cotton fields to allow calculation of the total (i.e., over the entire 2012-2013 season) chemical usage per unit area.

The distribution of the number of applications of *liquid Bollworm chemicals* is presented in Figure 5-13, and Figure 5-14 presents data on the total usage, over the 2012-2013 season, of these liquid Bollworm chemicals per hectare. Finally in this series, Figure 5-15 presents a scatterplot of *liquid Bollworm chemical per ha vs yield.* The average chemical usage for these farmers over the entire 2012-2013 season was 1789 ml/ha. The data for all of these figures were taken from the DZL (Central) Zambia survey.

Figure 5-13: Number of applications of liquid Bollworm chemicals, 2012-2013, DZL (Central) Zambia

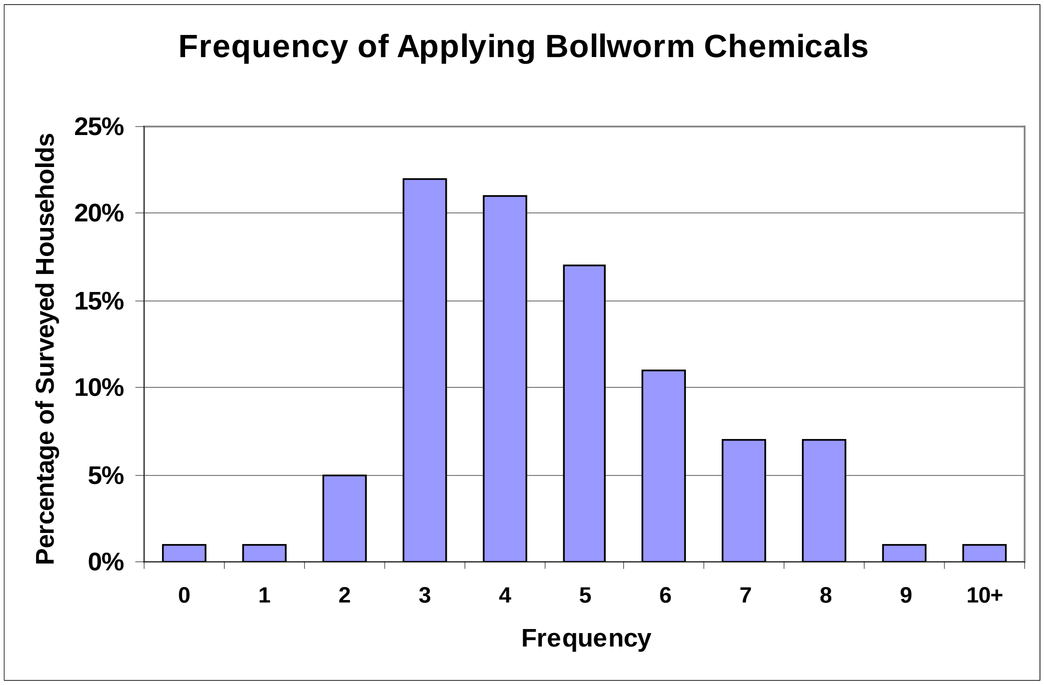


Figure 5-14: Usage of liquid Bollworm chemicals per hectare, 2012-2013, DZL (Central) Zambia

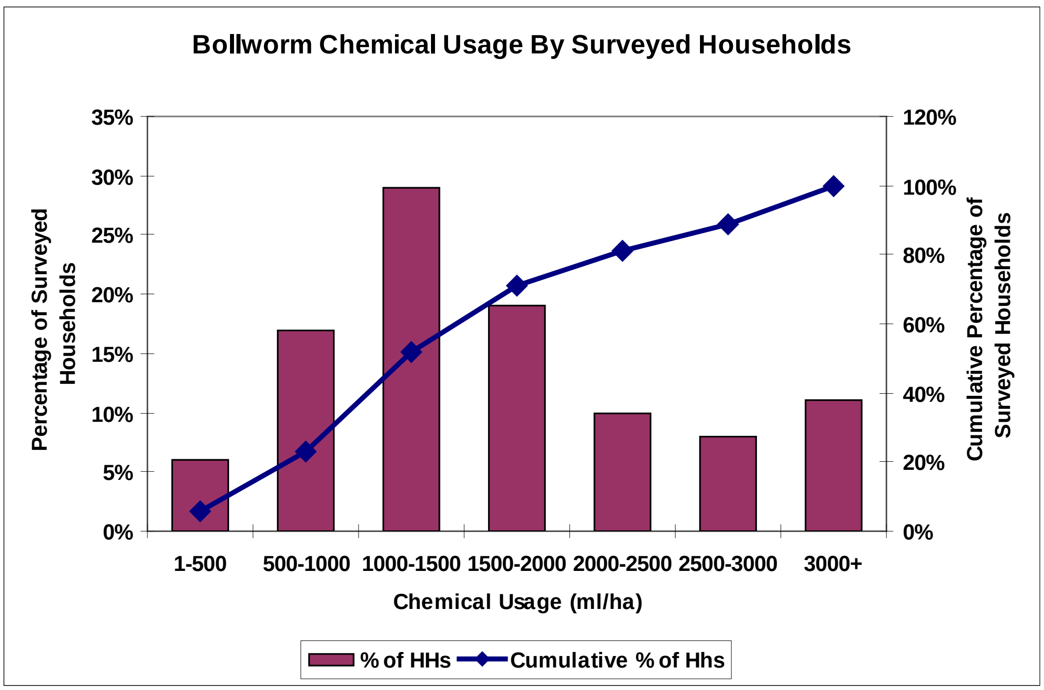


Figure 5-15: *Cotton yield* vs *liquid Bollworm chemical per ha*, DZL (Central) Zambia

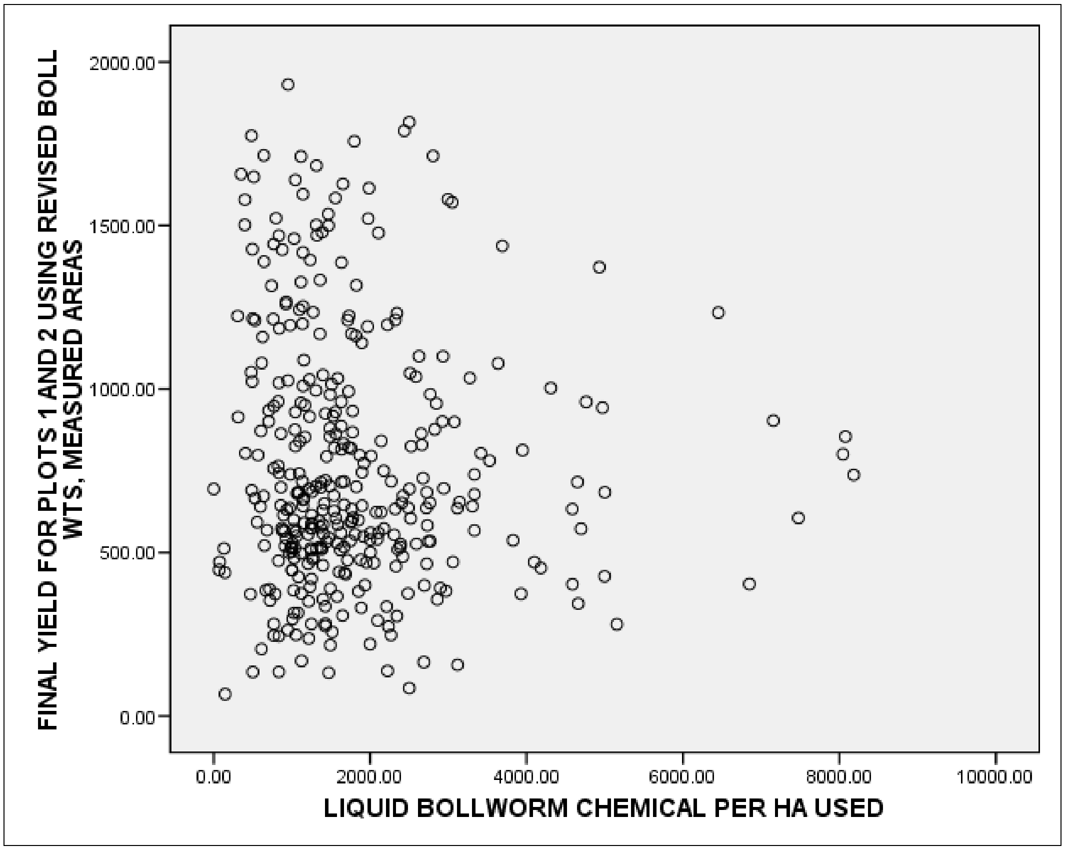


Figure 5-15 shows that, up to about 2500 ml/ha, cotton yield increases with increasing chemical use, but after about 2500 ml/ha, yield *decreases* with increasing chemical use. This is borne out by the Bivariate Coefficient of Correlation for these two variables which, at -0.06 shows an overall slightly inverse relationship; however, this relationship is not statistically significant (significance = 0.296). The available data do not indicate whether this decrease is due to overuse of the bollworm chemical or to other factors such as poor farming practices.

Other results once again show conclusively that the farmers do not know the sizes of their cotton plots with any accuracy. Figure 5-16 shows the distribution of values of the *ratio-of measured-to estimated cotton plot sizes.* This ratio was calculated for each surveyed Alliance Zambia farmer by dividing the size (area) of the primary cotton plot as determined by measurement with a GPS by what the farmer *thinks* the area of the plot is. (Only the 85% of farmers that provided these estimates are included in this figure.) Thus, if the farmer *overestimates* the plot size, the ratio will be less than 1.0; if (s)he *underestimates* the plot size, the ratio will be greater than 1.0. If, the farmer's estimate of the plot size is the same as the measured size, then this ratio will be exactly 1.0.

Figure 5-16: Ratio of measured-to-estimated sizes of the surveyed households' primary cotton plot (Alliance Zambia)

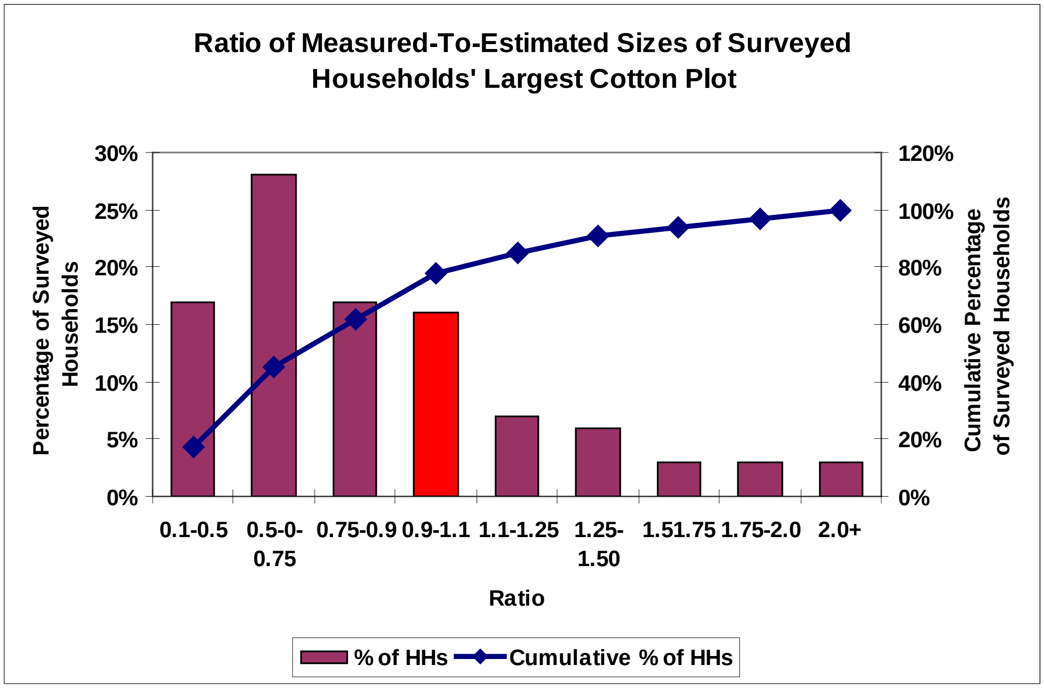
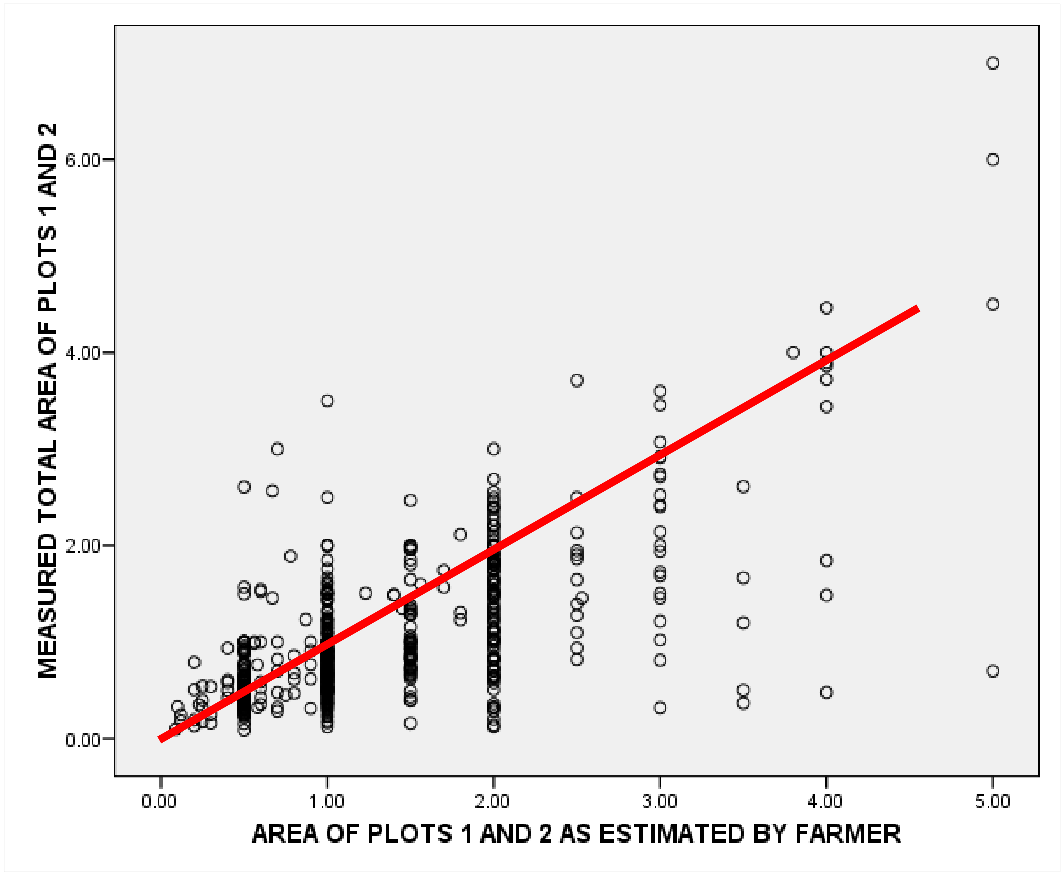


Figure 5-16 shows that most of the farmers *overestimate* their plot sizes, perhaps to get extra inputs on credit from the ginneries. In fact, almost half (45%) of these farmers overestimate their plot sizes by 11%-100% (ratio varies from 0.5 – 0.9). This figure shows that only about 1/6 (16%) of the farmers estimated their plot sizes to within ± 10% of the correct, measured value.

The incorrect farmer estimates of the sizes of their cotton plots have significant implications for the inputs provided by the ginneries (on credit) to these farmers. If the farmers overestimate their plot sizes and these estimates are the basis on which cotton seed is provided, then a great many farmers are receiving more seed than is needed. If this seed is provided on credit and if there is a high default rate on this credit, then the ginneries are losing even more money from these credit defaults than necessary.

These assertions are confirmed by a simple analysis of these data from Alliance Zambia. Figure 5-17 shows a scatterplot of the farmers' estimates of the total size of (up to) two cotton plots *vs* the measured sizes of these plots. The red diagonal line represents the case where the farmer's estimate equals the measurements. Note that most of the data points are below this line, indicating higher estimates than the true measured sizes for virtually all true plot sizes. In other words, these farmers overestimate plot areas for all plot sizes, not just for larger plots.

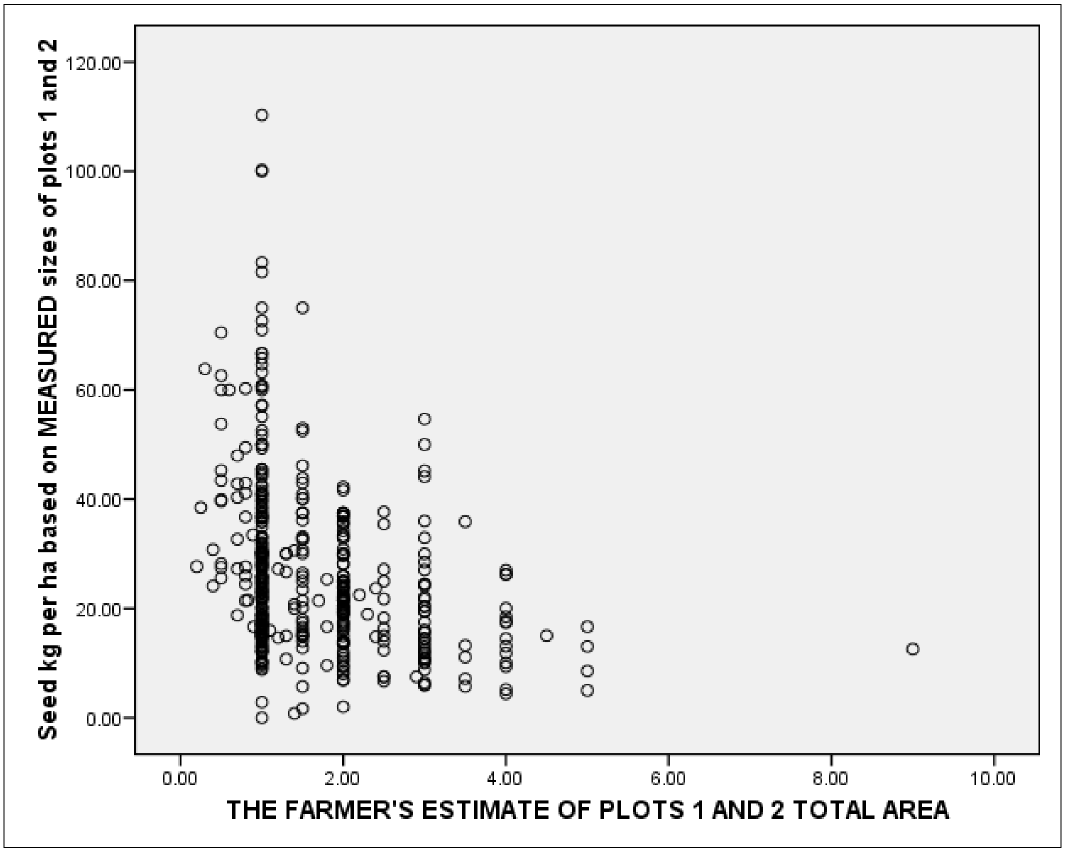
Figure 5-17: Measured vs estimated cotton plot sizes, Alliance Zambia



Alliance Zambia distributed seeds on the basis of these *estimated* plot sizes since, at the time they were not using GPS devices to measure plot sizes. Their policy is to distribute seed at the rate of 15 kg per ha of cotton field, but because of the errors in these estimates, the true amount of seed distributed per hectare varied from less than 5 kg/ha up to about 100 kg/ha. Only 43% of the farmers got within ± 50% of 15 kg/ha (7.5 kg/ha - 22.5 kg/ha) and 21% got between two and three times the intended amount (30 kg/ha – 45 kg/ha). Roughly 14% of the farmers got more than three times the intended amount.

These results are not unique to Alliance Zambia; another COMPACI Sub-Grantee in another country (Cargill Zimbabwe) experienced similar results as shown in Figure 5-18. This figure plots the farmers' estimates of the total area of (up to) two cotton plots against the amount of seed they received per *measured hectare* of (up to) two cotton plots. This figure shows that for a farmer's estimate of one hectare, the amount of seed given actually varied from about 10 kg/ha up to over 70 kg/ha.

Figure 5-18: Estimated plot sizes *vs* kg of seed per hectare give, Cargill Zimbabwe



In fairness, some of these results are due to the fact that seed is given in packets sized for 1 ha plots, and so cannot be given in 0.5 ha amounts. Thus, it is reasonable that farmers that have plots of "roughly" 1 ha be given a 1 ha seed packet. However, that alone does not explain the very large variance in seed given per ha shown in this figure; that is more likely a result of deliberate mis-statement of plot sizes in order to get more inputs, some of which may be intended for diversion to other crops or resale to other farmers.

5.3 Comparative Results

This section of the report presents some of the results from the various pilot yield surveys, but compares these same results across the different surveys. Note that because of the small number of farmers (240) and the statistically poor survey results from the Cargill Zambia (due to the inaccurate farmer list used to develop the sample), the results from this survey were not analyzed and are therefore not presented in this section.

Table 5.1 presents data from each survey on the average measured cotton area, the average estimated percentage of total cultivated land planted in cotton, and the average size of the total cultivated area. For these calculations, only the (up to) two largest cotton plots were used as these were the only cotton plots that were measured. This is a small restriction which includes all of the cotton areas for almost all of the farmers (for example, in the Zimbabwe Alliance survey, 789 (93.3%) of the farmers had one plot, 27 (3.3%0 had two cotton plots, and only one (0.1%) had more than one plot (28 farmers – 3.3% did not answer the question)).

The farmers were also asked "If the entire area on which you cultivate ALL crops were divided into 10 parts, how many of those parts did you use for COTTON in the last cropping season?" This, of course, is just the farmer's estimate, to the nearest 10%, of how much of his/her cultivated area was planted in cotton. From the total measured cotton area and this estimated percentage, the total area cultivated by that farmer can be estimated. The average values of these parameters for each of the surveys are presented in Table 5-1.

Table 5-1: Cotton areas and estimated total cultivated areas

Country	Cotton Company	Average *Measured* Cotton Area (2012-2013 Season) (Ha)	Average Percentage of Total Cultivated Area Planted in Cotton (2012-2013 Season)	Average Estimated Total Cultivated Area (2012-2013 Season) (Ha)
Zimbabwe	Alliance	1.09	40%	4.01
	Cargill	1.24	35%	5.07
Zambia	Alliance	1.00	39%	4.28
	Dunavant (Central)	1.07	34%	4.89
	Dunavant (Eastern)	0.78	38%	2.12

A comparison of average yields for the farmers using different land preparation methods is presented in Table 5-2; this table shows, for each of the five surveys compared here, the average yield on their *primary* cotton plot for the different land preparation methods used. Because the numbers and percentages of surveyed farmers using each land preparation method in each country varied significantly, these numbers and percentages of farmers using the different methods are also included in this table.

So, for example, this table shows that for the Alliance Zimbabwe survey, there were 78 farmers (10% of the sample farmers) that primarily used ripping for land preparation, and that on average these farmers got a yield of 687 kg/ha. Note that the different cotton companies and surveys used different boll weights, making direct comparisons of the yields across the different surveys highly problematic. However, within any one survey the different yields gained with the different land preparation methods can be compared.

Table 5-2: Average yields for different land preparation methods

Country	Cotton Company	Average Yields (kg/ha) and Land Preparation Method			
		Ripping	Ploughing	Potholing	Ridging
Zimbabwe	Alliance	687	686	832	N/A
		78 (10%)	735 (90%)	2 (0%)	0 (0%)
	Cargill	651	668	616	N/A
		81 (9%)	780 (90%)	8 (1%)	0 (0%)
Zambia	Alliance	606	537	415	753
		64 (8%)	699 (87%)	35 (4%)	1 (0%)
	Dunavant (Central)	648	694	672	661
		73 (18%)	309 (78%)	10 (3%)	3 (1%)
	Dunavant (Eastern)	719	636	1432	879
		14 (5%)	105 (38%)	11 (4%)	148 (53%)*

* This figure seems to be incorrect as it is a clear outlier compared to other regions of Zambia and should be lower than the figure for ploughing, because ridging is mostly done after ploughing

The numbers and percentages of farmers that received *any* training related to cotton farming and, more specifically, training on scouting and threshold spraying can also be compared. These data are presented in Table 5-3. In this case, since not all of the farmers answered all, or even the same, questions, the calculated percentages are based on the numbers of *responding* farmers, i.e., only those farmers that answered the applicable question(s).

Table 5-3: Numbers and percentages of farmers that received training

Country	Cotton Company	Numbers and Percentages of *Responding Surveyed* Farmers That Received Any Training Related to Cotton Farming	Numbers and Percentages of *Responding Surveyed* Farmers That Received Training Related to *Scouting and Threshold Spraying*
Zimbabwe	Alliance	458 (54%)	474 (58%)
	Cargill	834 (95%	816 (94%)
Zambia	Alliance	441 (55%)	427 (53%)
	Dunavant (Central)	287 (72%)	286 (72%)
	Dunavant (Eastern)	214 (77%)	204 (73%)

Table 5-4 is similar to Table 5-3, except here the gender of those farmers trained is considered. In this table, the percentages given for men and for women that received training are *the percentages of men and women that were interviewed; this is not the same as the percentage of people that reported being trained that are men and women*. The percentages of female farmers that were interviewed in each of the surveys are also included in this table.

Table 5-4: Numbers and percentages of farmers that received training

Country	Cotton Company	Percentage of Surveyed Farmers That Are Female	Numbers and Percentages of *Responding Surveyed Male and Female* Farmers That Received Any Training Related to Cotton Farming		Numbers and Percentages of *Responding Surveyed Male and Female* Farmers That Received Training Related to *Scouting and Threshold Spraying*	
			Men	Women	Men	Women
Zimbabwe	Alliance	23%	59%	46%	61%	48%
	Cargill	23%	96%	95%	95%	92%
Zambia	Alliance	21%	56%	52%	54%	52%
	Dunavant (Central)	20%	75%	61%	75%	58%
	Dunavant (Eastern)	27%	81%	67%	77%	63%

The types of equipment used for spraying pesticides and the household members responsible for "most" of the spraying can also be compared across the different yield surveys. These data are presented in Table 5-5; again, the percentages presented in this table are based on the numbers of *responding* farmers only.

Table 5-5: Pesticides spraying equipment and intra-household responsibility for spraying

Country	Cotton Company	Type of Spraying Equipment		Primary Responsibility for Spraying				
		Knapsack Sprayer	Ulva and Sprayer	Farmer	Other Adult in HH	Child in HH	Hired Labor*	Other
Zimbabwe	Alliance	807 (99%)	6 (1%)	742 (91%	70 (9%)	2 (0%)	3 (0%)	0 (0%)
	Cargill	862 (99%)	4 (1%)	763 (88%)	99 (11%)	1 (0%)	3 (0%)	2 (0%)
Zambia	Alliance	741 (93%)	6 (1%)	661 (83%)	123 (15%)	4 (0%)	11 (1%)	0 (0%)
	Dunavant (Central)	357 (89%)	15 (4%)	332 (84%	52 (13%)	4 (1%)	4 (1%)	5 (1%)
	Dunavant (Eastern)	265 (95%)	2 (1%)	203 (73%)	70 (25%)	0 (0%)	5 (2%)	0 (0%)

* "Hired labor" includes service providers

Finally, a topic of great interest to both COMPACI and to CmiA is *primary school attendance rates.* Questions regarding the numbers of boys and girls in the household of primary school age (which varies by country) and the numbers of them regularly attending primary school were asked as part of all Yield Estimation Surveys. ("Regularly attending primary school" was defined for these surveys as *attending at least 3 weeks of every month when school is in session.*)

Table 5-6 presents the primary school attendance data from the 2010 Baseline Survey from Zambia's Eastern Province and all of the pilot yield surveys (except for Cargill Zambia whose complete dataset had not been received until quite late and which ultimately contained a small (240 households) non-representative sample of their farmers.)

Table 5-6: Percentages of primary school children attending primary school

Country	Cotton Company	Percentage of Boys Attending Primary School	Percentage of Girls Attending Primary School
Zimbabwe	Alliance	73%	80%
	Cargill	76%	83%
Zambia	Alliance	86%	87%
	Dunavant (Central)	88%	87%
	Dunavant (Eastern)	88%	94%
	2010 Baseline Survey (Eastern Province)	61%	64%

6.0 The Way Forward

This section discusses what needs to be done in 2014 and onward to improve the reliability and accuracy of the survey results and the utility to COMPACI and to the Sub-Grantees of these surveys. The challenges faced and their proposed solutions are presented in Section 6.1. A proposed methodology for transferring the survey technology to the COMPACI Sub-Grantees so they can independently conduct yield estimation surveys and studies is presented in Section 6.2.

6.1 Challenges and Proposed Corrective Actions

6.1.1 Boll Weights

The principal challenge to the success and accuracy of the Yield Estimation Survey methodology described in this report is the need for improvement in the accuracy of the boll weights provided by the ginneries. Recall from earlier discussion in this report that an error of X% in the assumed average boll weight for any cotton field results directly in an X% error in the estimates of both the field's production and yield apart from any other errors that may occur..

Part of the challenge here stems from the fact that boll weights can vary not only by the variety of cotton planted, but also by the agro-ecological conditions (soil, altitude, weather – rains) of the field, and by planting date. ("*Planting date*" is in this case a proxy for the time of planting relative to the onset of the rains.) While it is expected that the ginneries can and will, with more experience with these surveys, refine the weighing of bolls from different areas in order to produce more accurate *average* boll weights for an area that season, it may be very difficult for them to accommodate different planting dates in these measurements as is already done by Faso Coton in Burkina Faso.

Therefore, some of the planting date variations in boll weight will be resolved either by accessing data that may be available in research institutions or by limited experiments to determine, for a given Sub-Grantee and season, the effects of *planting date* on boll weight. This might be done by asking farmers when they planted their cotton (as is already done in the current yield surveys) and then weighing samples of boll weights and tracking average boll weights against the planting dates. Once a "library" of these data are built up, the effect of planting date should be able to be handled in the yield calculations by a lookup table that relates the percentage weight gain or loss according to delay after the onset of the rains in planting the cotton.

6.1.2 Enumerator Errors

Analysis of the collected survey data from several of the pilot yield estimation surveys make clear that there are some types of enumerator error that occur with some regularity. These include:

- Improper placement or omission of decimal points in entering areas and/or boll weights into the tablets
- Confusing questions and entering incorrect answers about *row separation distances* vs *station gap distances*
- Not asking farmers for more information when the answers to some questions are not credible or are unclear

All of these errors occur relatively infrequently and so did not cause any significant problems in the analysis. However, they are all preventable. The plans for 2014 include short refresher trainings for Zambia and Zimbabwe and longer, initial trainings where these surveys will be conducted for the first time. The training facilitator will take special care to remark on these errors and to emphasize the corrective actions that should be taken by the enumerators to prevent these from occurring.

Since, by design, the enumerators are all ginnery extension staff, they are already familiar with the agronomy and practices of cotton farming and so should be easily able to get more correct and accurate responses from the surveyed farmers and enter these responses correctly.

6.1.3 Refinement of the Yield Estimation Methodology

After reviewing the similar methodologies used by some of the West African COMPACI Sub-Grantees, there are several changes that will be made to the Yield Estimation Methodology used in the 2013 Pilot Surveys. These changes are designed to improve the accuracy of the estimation of the number of bolls in the field (which is used, together with the assumed average boll weight, to estimate the field's production.) Analysis of the pilot data is still ongoing, so other changes in addition to those described below may also be required. Note, however, that the fundamental methodology and the equations in Annex 1 used to implement this methodology are correct.

The principal changes planned at this time are:

- Increase the length of the sample strips in which the in-field measurements are made from 3 meters to 5 meters. This will result in a 67% larger sample of the field's plants being used in the estimation than was used in the Pilot Surveys.

- Provide training by COMPACI agronomists on how to select “random” sample strips in the surveyed cotton fields and how to recognize viable cotton bolls and to count them. This training would be one “module” in the overall survey training.

6.2 Technology Transfer to Ginneries

COMPACI Management expects that the Sub-Grantees will continue to implement these Yield Estimation Surveys, in some form, after the end of COMPACI. The Sub-Grantees already have the necessary technology – the tablets and the GPS devices, funded by COMPACI. However, they do not have the software licenses needed to implement the survey on the tablets and upload the collected data to some server from which the data would be retrieved and analyzed. Also, in general the cotton companies do not have the data analysis capability or software needed to produce the kinds of results presented in this report.

Therefore, there are a range of possible options available for the desired technology transfer, ranging from simple, limited capability models requiring little or no data analysis expertise or expenses for software licenses, software hosting on servers or data analysts, up to fairly sophisticated options that would more or less enable the ginneries to conduct these surveys as the pilot surveys were conducted and to have the data analysis done by external consultants.

Because this is, for most of the COMPACI Sub-Grantees, a new initiative with somewhat uncertain benefits to their operations and bottom line, it is likely that, for the near-term at least, they would prefer to minimize their outlays of money or manpower until they have gained more experience with these surveys and better understand and appreciate the benefits that these surveys can provide to them. Therefore, the *less-is-more* minimal option for this technology transfer is recommended here. This simpler, less demanding approach is discussed below.

The basic concept for this technology transfer model is to use an Excel spreadsheet to list the farmers to be surveyed (which can be any subset of farmers from any region, selected for any ginnery reason or purpose and need not be selected from some representative sample) and up to three geographical variables. These geographical variables could be Province, District, Ward/Village or they could be Ag Office, Shed, and Zone; in short they can be whatever the Sub-Grantee wants to use to describe the hierarchy of administrative survey area(s).

This Excel spreadsheet would also have pre-labeled columns for some of the in-field measurements and some of the results calculated and (currently) displayed by the tablets into which the results of these measurements would be entered, just as was done in the 2013 pilot surveys. As presently envisioned, this Excel spreadsheet would have pre-defined columns for *plot size, estimated production, estimated yield, average row separation, average station gap,* and *average number of bolls per plant*

The enumerators would measure the areas of the selected cotton fields with their GPS and, in the field, manually enter these areas into the Excel spreadsheet. The enumerators would then be guided by the tablets to select the 5-8 sample strips of cotton rows (each strip would now be 5 meters rather than the 3 meters used this past season) as was done in 2013. In each sample strip the enumerator would be guided, as was done in the Pilot Surveys, to measure the left and right row separation distances, and to count the number of plants and the number of viable bolls; these data would all be entered into the tablets, again as was done in 2013.

When the complete set of farmers' fields had been thus surveyed, the data from the manual Excel spreadsheet would be entered into a computerized Excel file provided by NORC that has a built in pivot table function. This pivot table function would automatically take the collected farmer field-level data, and for each different value of the geographical variables, the pivot table would display the average value of *plot size, estimated production, estimated yield, average row separation, average station gap,* and *average number of bolls per plant.*

Thus, if the geographical variables were *province, district, and village*, then the pivot table would automatically display the calculated average values of *plot size, estimated production, estimated yield, average row separation, average station gap,* and *average number of bolls per plant* for each village, district, and province that occurs in the collected data.

This technology transfer option would provide the ginneries with the ability to do yield estimation surveys in any area and of any group of farmers without any external support. As the ginneries gain experience with this tool and realize they would like something more sophisticated, they could explore options for gaining this increased survey capacity by either increasing their internal capacity or by engaging external support.

Annex 1: Details of the In-Field Sampling Approach to Yield Estimation

Method for Estimating Cotton Plant Population, Production, and Yield of a Single Plot

1. Method concept

This method works as follows

1. Measure the area of the plot with a GPS
2. Determine the average row separation between rows of cotton
3. Use this calculated separation distance and measured area to calculate the length of the equivalent one-row cotton plot in meters (this step converts the actual shape of the plot to an assumed rectangle of the width = average row separation distance and calculated length that has the same area as the actual plot)
4. The total *linear meters of cotton* in the plot (plot length based on assumed rectangular plot shape)
5. From a number of 3 meter row sections sampled from the plot, determine the *average number of cotton plants and cotton bolls per 3 meter section of a row*
6. Estimate total numbers of cotton plants and the number of cotton bolls in the entire plot by scaling up from the average for 3 meters to the total linear meters of cotton in plot as calculated above
7. Multiply estimated number of cotton bolls in the plot by average boll weight (grams) provided by the cotton company and divide by 1000 to calculate *estimated total cotton production (kg) for the plot.*
8. Calculate estimated yield of the plot by dividing estimated production (kg) by the measured area (ha). **IF MULTIPLE PLOTS FOR ONE FARMER ARE INVOLVED, THIS CALCULATION SHOULD BE REPLACED BY ESTIMATED PRODUCTION OF (UP TO) 2 LARGEST PLOTS DIVIDED BY THE MEASURED AREA OF ALL PLOTS.** ***Note that this does not include production of plots 3, 4 etc. This was done too reduce possibly excessive field time per farmer this year***
9. Calculate *average gap spacing* and *average number of bolls per cotton plant*

2. Data: This method uses the following data for each sampled plot:

1. The data from each of the 5-8 sampling stations sampled for the plot. Total sampling stations in that plot = $\mathbf{N}$
   a. Distance (cm) to row to the LEFT of sampled row i: $\mathbf{R_{Li}}$
   b. Distance (cm) to row to the RIGHT of sampled row i: $\mathbf{R_{Ri}}$
   c. Number of cotton PLANTS along the sample 3 meter section of the sample row i: $\mathbf{P_i}$
   d. Number of cotton BOLLS along the sample 3 meter section of the sample row i: $\mathbf{B_i}$
2. The area of the plot (ha) $\mathbf{A}$
3. The average boll weight (grams) (provided by cotton company for that area): $\mathbf{G}$

3. Calculations

1. Calculate the average row separation with the average left and right row separation distances from all sampling stations in the plot (**N** sample rows, 2N measurements in all)

$$R_{cm} = \left(\sum_{i=1}^{N} (R_{Li} + R_{Ri}) \right) / 2N$$

2. The assumed width of the equivalent rectangular cotton plot one row wide (in meters)

$$W = R_{cm} / 100$$

3. Calculate the *total linear meters of cotton* in the plot, **L**, assuming that the plot is rectangular with width **W**

$$L = (A_j \times 10000) / W$$ **where A_j is the area of that plot only (j = 1, 2)**

4. Calculate estimated **Plant Population** (total number of cotton plants), **P**, of plot

$$P = \left(\left(\sum_{i=1}^{N} P_i\right) / N \right) \times (L / 3)$$

5. Calculate estimated **Number of Bolls** (total number of cotton bolls), **B**, of plot

$$B = \left(\left(\sum_{i=1}^{N} (B_i) \right) / N \right) \times (L / 3)$$

6. Calculate the total estimated production (kg), **KG,** of the plot

$$KG = (B \times G) / 1000$$

7. Calculate estimated yield (kg/ha), **Y** of the plot for that farmer

$$Y = KG / A$$

8. Calculate average number of cotton bolls per plant, **BP** in the plot

$$BP = \left(\sum_{i=1}^{N} (B_i / P_i) \right) / N$$

9. Calculate average Gap Spacing (cm) between cotton plants in the same row, **S**

$$S = \left(\sum_{i=1}^{N} (300 / (P_i - 1)) \right) / N$$

Annex 2: Sample Questionnaire Used in 2013 Pilot Surveys in Zambia

2013 ZAMBIA COTTON YIELD SURVEY - DATA COLLECTION FORM (GPS)

1.1	FARMER ID NUMBER	|__|__ | - |__|__|__|__|__|__| |__| - |__|__|		**SURVEY MANAGEMENT**	
1.2	**FARMER NAME (Last, first, middle)**	|____________||____________||____________|		ENUMERATOR'S NAME	|______________________|
1.3	FARMER GENDER	|__| MALE = 1; FEMALE = 2; DON'T KNOW = -1; REFUSED = -2		ENUMERATOR'S CODE	|__|__|__|
1.4	PROVINCE	|______________________________________|		DATE OF INTERVIEW	|__|__|__|__|__|__|
1.5	DISTRICT	|______________________________________|			
1.6	DEPOT/AGRIC. OFFICE	|______________________________________|		SUPERVISOR'S CODE	|__|__|__|
1.7	**VILLAGE/DISTRIBUTOR**	|______________________________________|			

INTRODUCTION AND CONSENT

Hello and thank you for talking to us. We are from[COTTON COMPANY]. We are working on the Competitive African Cotton Initiative (COMPACI) Program that is intended to increase the income of cotton farming households. This program is funded by DEG, a German development bank.

As part of this program, we are collecting data on the yield of cotton plots for cotton farmers in Zambia. We are asking these questions and seeking your answers in order to understand the impacts of this program on farming households. This information is important to know as it will tell us how to improve the success of the program to support cotton farmers in local communities. We would like you to allow us to measure the size of the plot that you used to grow cotton in the 2012/2013 season and to answer some questions about growing and harvesting cotton in 2012/2013. We expect the total time required for the interview will be at most 45 minutes, plus the time needed for assessing your cotton field. The specific purpose of this survey is to obtain accurate information on the actual yields of cotton farmers and compare them with estimates produced by cotton companies and the Ministry of Agriculture. This survey may be repeated in 2014 and following years and you may be asked to participate again so that we can measure the changes in yield among cotton farmers.

The information that we collect today will be compared to data from [COTTON COMPANY]with whom you worked in 2012/2013; the responses you provide today, including your name and location, may be shared with the cotton company and DEG. The information that you provide will be kept until at least 2015 for the purposes of preparing reports on the changes among cotton growers. Your participation is completely voluntary. You are free to not answer any question with which you are not comfortable, and you may stop participating in the survey at any time. If you have any questions about the survey, you may contact **[CONTACT INFORMATION]** at the telephone number or address on this card.

GIVE CARD TO RESPONDENT.

1.8	Do you wish to participate in this survey?	|__|	1 YES 2 NO *DK/NA/REFUSED NOT ALLOWED*

I would like to start with some general questions.

2	**GENERAL QUESTIONS**		
2.1	Besides cotton, what are the most important crops, and how many hectares did you grow of each? SELECT UP TO THREE	|__|(___________) |__|(___________) |__|(___________)	1 MAIZE 2 GROUNDNUTS 3 SUNFLOWER 4 SORGHUM 5 BEANS 6 COWPEAS 7 MILLET 8 TOBACCO 9 SOYBEAN 10 OTHER 1 SPECIFY (_____) 11 OTHER 2 SPECIFY (_____) 12 OTHER 3 SPECIFY (_____) -1 DON'T KNOW -2 REFUSED
2.2	How many years has someone in this household grown cotton? IF LESS THAN ONE YEAR ENTER ZERO -1 = DON'T KNOW -2 = REFUSED	|__|__|	NUMBER OF YEARS
2.3	How many separate training sessions, meetings, or demonstrations in basic agricultural techniques has the principal cotton farmer in this household attended in the last 12 months (May 2012-April 2013)? IF NONE, ENTER 0 -1 = DON'T KNOW -2 = REFUSED	|__|__|	NUMBER OF TRAINING SESSIONS
2.4	How many separate cotton plots did you have last season (2012-2013)? -1 = DON'T KNOW -2 = REFUSED	|__|	NUMBER OF COTTON PLOTS (ZERO IS NOT A VALID RESPONSE)
2.5	If the entire area on which you cultivate ALL crops were divided into 10 parts, how many of those parts did you use for COTTON in the last cropping season (2012-2013)? -1 = DON'T KNOW -2 = REFUSED	|__|__|	NUMBER OF PARTS OF TOTAL CULTIVATED AREA USED FOR COTTON (ZERO IS NOT A VALID RESPONSE)

FARMING PRACTICES

Now, I will ask you some questions about some of your farming practices for the 2012/2013 season.

3	**SOIL FERTILITY - CROP ROTATION**		
3.1	Do you practice crop rotation on ALL of your cotton field(s)?	|__|	1 YES ***GO TO Q3.1b*** 2 SOME PART3 NO GO TO 3.2 -1 DON'T KNOW -2 REFUSED
3.1a	If the entire area on which you cultivated *COTTON* in 2012-2013 were divided into 10 parts, on how many of those parts did you *USE A ROTATION CROP* in the previous cropping season (2011-2012)? -1 = DON'T KNOW -2 = REFUSED	|__|__|	NUMBER OF PARTS OF TOTAL COTTON AREA WHERE CROP ROTATION WAS PRACTICED
3.1b	Which are the (up to) 3 primary crops that you use for rotation with cotton?	|__|(____________) |__|(____________) |__|(____________)	1 MAIZE 2 GROUNDNUTS 3 SUNFLOWER 4 SORGHUM 5 BEANS 6 COWPEAS 7 MILLET 8 TOBACCO 9 SOYBEAN 9 OTHER 1 SPECIFY (____) 10 OTHER 2 SPECIFY (____) 11 OTHER 3 SPECIFY (______) -1 DON'T KNOW -2 REFUSED
3.2	What other practices do you use to improve soil fertility? MARK ALL THAT APPLY	|__| |__| |__| |__| |__|	1 RESIDUE MANAGEMENT / MULCHING 2 COMPOST 3 MANURE 4 FERTILIZER 5 LIME 6 OTHER 1 SPECIFY (____) 7 OTHER 2 SPECIFY(_____) -1 DON'T KNOW -2 REFUSED

4	SOIL FERTILITY – USE OF MULCHING / COMPOST / MANURE / LEGUMINOUS CROPS /FERTILIZER / LIME		
4.1	If the entire area on which you cultivate *ALL* crops were divided into 10 parts, on how many of those parts did you use *MULCHING* during the last cropping season (2012-2013)? IF NONE, ENTER 0 -1 = DON'T KNOW -2 = REFUSED	|__|__|	NUMBER OF PARTS OF TOTAL CULTIVATED AREA ON WHICH *MULCHING* WAS USED
4.2	If the entire area on which you cultivate *ALL* crops were divided into 10 parts, on how many of those parts did you apply *COMPOST* during the last cropping season (2012-2013)? IF NONE, ENTER 0 -1 = DON'T KNOW -2 = REFUSED	|__|__|	NUMBER OF PARTS OF TOTAL CULTIVATED AREA ON WHICH *COMPOST* WAS USED
4.3	If the entire area on which you cultivate *ALL* crops were divided into 10 parts, on how many of those parts did you apply *MANURE* during the last cropping season (2012-2013)? IF NONE, ENTER 0 -1 = DON'T KNOW -2 = REFUSED	|__|__|	NUMBER OF PARTS OF TOTAL CULTIVATED AREA ON WHICH *MANURE* WAS USED
4.4	If the entire area on which you cultivate *ALL* crops were divided into 10 parts, on how many parts did you grow *LEGUMINOUS CROPS* (beans, groundnuts, etc.) last cropping season (2012-2013)? IF NONE, ENTER 0 -1 = DON'T KNOW -2 = REFUSED	|__|__|	NUMBER OF PARTS OF TOTAL CULTIVATED AREA ON WHICH *LEGUMINOUS CROPS* WERE GROWN
4.5	If the entire area on which you cultivate *ALL* crops were divided into 10 parts, on how many parts did you use *COMPOUND FERTILIZER* in the last cropping season (2012-2013)? IF NONE, ENTER 0 -1 = DON'T KNOW -2 = REFUSED IF 0, SKIP TO 4.7	|__|__|	NUMBER OF PARTS OF TOTAL CULTIVATED AREA ON WHICH *COMPOUND FERTILIZER* WAS USED

4.6	How many kg of *COMPOUND FERTILIZER* did you use in the last cropping season (2012-2013)? -1 = DON'T KNOW -2 = REFUSED	|__|__|__|__|	KG OF *COMPOUND FERTILIZER* USED
4.7	If the entire area on which you cultivate *ALL* crops were divided into 10 parts, on how many parts did you use *NITROGEN TOP DRESSING* in the last cropping season (2012-2013)? IF NONE, ENTER 0 -1 = DON'T KNOW -2 = REFUSED IF 0, SKIP TO 4.9	|__|__|	NUMBER OF PARTS OF TOTAL CULTIVATED AREA ON WHICH *NITROGEN TOP DRESSING* WAS USED
4.8	How many kg of *NITROGEN TOP DRESSING* did you use in the last cropping season (2012-2013)?	|__|__|__|__|	KG OF *NITROGEN TOP DRESSING* USED
4.9	If the entire area on which you cultivate *ALL* crops were divided into 10 parts, on how many of those parts did you apply *LIME* during the last cropping season (2012-2013)? IF NONE, ENTER 0 -1 = DON'T KNOW -2 = REFUSED IF 0, SKIP TO SECTION 5	|__|__|	NUMBER OF PARTS OF TOTAL CULTIVATED AREA ON WHICH *LIME* WAS USED
4.10	How many kg of *LIME* did you use in the last cropping season (2012-2013)? -1 = DON'T KNOW -2 = REFUSED	|__|__|__|__|	KG OF *LIME* USED

COTTON FARMING PRACTICES

Now I will ask you about various farming practices particular to **COTTON**.

5	**LAND PREPARATION**		
5.1	What was the most important method for doing the land preparing of your cotton fields?	|__|	1 PLOUGHING 2 RIPPING 3 POTHOLING 4 OTHER SPECIFY(____) -1 DON'T KNOW -2 REFUSED
5.2	How was the land preparation mostly done?	|__|	1 OWN MANUAL LABOUR 2 HIRED MANUAL LABOUR 3 OWN OXEN 4 OWN TRACTOR 5 OXEN TILLAGE SERVICE PROVIDER 6 TRACTOR TILLAGE SERVICE PROVIDER -1 DON'T KNOW -2 REFUSED
5.3	Which month did you do most of the land preparation for your cotton fields?	|__|__|	1 JAN 2 FEB 3 MAR 4 APR 5 MAY 6 JUN 7 JUL 8 AUG 9 SEP 10 OCT 11 NOV 12 DEC -1 DON'T KNOW -2 REFUSED
5.4	If your entire cotton area were divided into 10 parts, how many of those parts did you use *RIPPING* in the last cropping season (2012-2013)? IF NONE, ENTER 0 -1 = DON'T KNOW -2 = REFUSED	|__|__|	

6	PLANTING		
6.1	In which month were the first good planting rains in your area?	|__|__|	1 JAN 2 FEB 3 MAR 4 APR 5 MAY 6 JUN 7 JUL 8 AUG 9 SEP 10 OCT 11 NOV 12 DEC -1 DON'T KNOW -2 REFUSED
6.2	In which part of the month was this?	|__|	1 BEGINNING 2 MIDDLE 3 END -1 DON'T KNOW -2 REFUSED
6.3	How many days after the first good planting rains did you plant your cotton?	|__|	1 0-7 DAYS 2 8-14 DAYS 3 15-21 DAYS 4 22-28 DAYS 5 MORE THAN 28 DAYS -1 DON'T KNOW -2 REFUSED
6.4	In what month did you plant your cotton?	|__|__|	1 JAN 2 FEB 3 MAR 4 APR 5 MAY 6 JUN 7 JUL 8 AUG 9 SEP 10 OCT 11 NOV 12 DEC -1 DON'T KNOW -2 REFUSED
6.5	In what part of the month did you plant your cotton?	|__|	1 BEGINNING 2 MIDDLE 3 END -1 DON'T KNOW -2 REFUSED
6.6	Did you have to re-plant your cotton?	|__|	1 YES 2 NO - GO TO **Q7** -1 DON'T KNOW -2 REFUSED
6.7	In what month did you re-plant your cotton?	|__|__|	1 JAN 2 FEB 3 MAR 4 APR 5 MAY 6 JUN 7 JUL 8 AUG 9 SEP 10 OCT 11 NOV 12 DEC -1 DON'T KNOW -2 REFUSED
6.8	In what part of the month did you re-plant your cotton?	|__|	1 BEGINNING 2 MIDDLE 3 END -1 DON'T KNOW -2 REFUSED

7	PLANT POPULATION		
7.1	How far apart are the rows of cotton plants? -1 = DON'T KNOW -2 = REFUSED	|__|__|__|	DISTANCE IN CM
7.2	How far apart were the planting stations for your cotton? -1 = DON'T KNOW -2 = REFUSED	|__|__|__|	DISTANCE IN CM
7.3	How many seeds did you plant per station? -1 = DON'T KNOW -2 = REFUSED	|__|	NUMBER OF SEEDS
7.4	Did you do any gap filling?	|__|	1 YES 2 NO ***GO TO Q 7.6*** -1 DON'T KNOW -2 REFUSED
7.5	How many days after first crop emergence did you do gap filling? IF ON THE SAME DAY, ENTER 0 -1 = DON'T KNOW -2 = REFUSED	|__|__|	
7.6	Did you thin your cotton plants (remove extra cotton plants from the same hole)?	|__|	1 YES 2 NO ***GO TO Q 8*** -1 DON'T KNOW -2 REFUSED
7.7	How many days after plant emergence did you thin? IF ON THE SAME DAY, ENTER 0 -1 = DON'T KNOW -2 = REFUSED	|__|__|	NUMBER OF DAYS
7.8	How many plants did you leave per station? -1 = DON'T KNOW -2 = REFUSED	|__|	NUMBER OF PLANTS

8	**WEED CONTROL**		
8.1	How did you do most of your weed control?	|__|	1 MANUAL - GO TO Q **8.5** 2 MECHANICAL – GO TO Q **8.5** 3 HERBICIDES 4 COMBINATION OF ABOVE -1 DON'T KNOW (GO TO SECTION 9) -2 REFUSED (GO TO SECTION 9)
8.2	Which herbicides did you use (up to 2 answers)? Names of herbicides used in Zambia to be provided by companies	|__| |__|	1 2 3 6 OTHER SPECIFY (____) -1 DON'T KNOW -2 REFUSED
8.3	When did you apply these herbicides?	|__|	1 PRE-EMERGENCE 2 POST-EMERGENCE 3 BOTH 4 OTHER SPECIFY (____) -1 DON'T KNOW -2 REFUSED
8.4	Were the herbicides effective in controlling weeds?	|__|	1 YES 2 PARTLY 3 NO -1 DON'T KNOW -2 REFUSED
8.5	How many times did you manually or mechanically weed your main cotton plot? -1 DON'T KNOW -2 REFUSED	|__|	NUMBER OF TIMES
8.6	How many weeks after plant emergence was the FIRST manual or mechanical weeding? -1 DON'T KNOW -2 REFUSED	|__|__|	WEEKS

9	PESTICIDE USE ON COTTON		
9.1	What do you use for applying your pesticides?	|__|	1 KNAPSACK SPRAYER 2 ULVA+ SPRAYER -1 DON'T KNOW -2 REFUSED
9.2	Who does most of the spraying?	|__|	1 FARMER HIMSELF/HERSELF 2 OTHER ADULT FROM HOUSEHOLD 3 CHILD (UNDER 18) FROM HOUSEHOLD 4 SPRAY SERVICE PROVIDER 5 OTHER -1 DON'T KNOW -2 REFUSED
9.3a	Did you receive training in scouting and threshold spraying?	|__|	1 YES 2 NO - GO TO ***Q 9.4a*** -1 DON'T KNOW -2 REFUSED
9.3b	If you used scouting, how many times did you scout? ENTER 0 FOR NONE -1 DON'T KNOW -2 REFUSED	|__|__|	
9.3c	How many times after scouting did you decide to spray less than 3 days after the scouting? IF NONE, ENTER 0 -1 DON'T KNOW -2 REFUSED	|__|__|	
9.4a	How many times did you apply **BOLLWORM** chemicals? ENTER 0 FOR NONE -1 DON'T KNOW -2 REFUSED	|__|__|	*IF RESPONSE = 0,* ***GO TO Q 9.5a***
9.4b	Did you apply the BOLLWORM chemicals at regular intervals or only when you detected pests on your plants (scouting)?	|__|	1 REGULAR INTERVALS 2 SCOUTING -1 DON'T KNOW -2 REFUSED
9.4c	How many **bottles** of BOLLWORM chemicals did you use in total? IF NONE, ENTER 0 -1 DON'T KNOW -2 REFUSED	|__|__|	
9.4d	How many ml is there in a **bottle** of BOLLWORM chemical? -1 DON'T KNOW -2 REFUSED	|__|__|__|	

<table>
<tr><td>9.4e</td><td>How many sachets of BOLLWORM chemicals did you use in total?

IF NONE, ENTER 0

-1 DON’T KNOW
-2 REFUSED</td><td>|__|__|</td><td></td></tr>
<tr><td>9.4f</td><td>How many grams are there in a sachet of BOLLWORM chemical?

-1 DON’T KNOW
-2 REFUSED</td><td>|__|__|__|</td><td></td></tr>
<tr><td>9.5a</td><td>How many times did you apply APHID chemicals to your cotton plants during the last cropping (2012-2013) season? ENTER 0 FOR NONE</td><td>|__|__|</td><td>IF RESPONSE = 0, GO TO Q 9.6a</td></tr>
<tr><td>9.5b</td><td>Did you apply the APHID chemicals at regular intervals or only when you detected pests on your plants (scouting)?</td><td>|__|</td><td>1 REGULAR INTERVALS
2 SCOUTING
-1 DON’T KNOW
-2 REFUSED</td></tr>
<tr><td>9.5c</td><td>How many bottles of APHID chemicals did you use in total?

IF NONE, ENTER 0

-1 DON’T KNOW
-2 REFUSED</td><td>|__|</td><td></td></tr>
<tr><td>9.5d</td><td>How many ml is there in a bottle of APHID chemical?

-1 DON’T KNOW
-2 REFUSED</td><td>|__|__|__|</td><td></td></tr>
<tr><td>9.5e</td><td>How many sachets of APHID chemicals did you use in total?

IF NONE, ENTER 0

-1 DON’T KNOW
-2 REFUSED</td><td>|__|__|</td><td></td></tr>
<tr><td>9.5f</td><td>How many grams are there in a sachet of APHID chemical?

-1 DON’T KNOW
-2 REFUSED</td><td>|__|__|__|</td><td></td></tr>
<tr><td>9.6a</td><td>How many times did you apply STAINER chemicals to your cotton plants during the last cropping (2012-2013) season? ENTER 0 FOR NONE

-1 DON’T KNOW
-2 REFUSED</td><td>|__|__|</td><td></td></tr>
</table>

9.6b	Did you apply the STAINER chemicals at regular intervals or only when you detected pests on your plants (scouting)?	|__|	1 REGULAR INTERVALS 2 SCOUTING -1 DON'T KNOW -2 REFUSED
9.6c	How many **bottles** of STAINER chemicals did you use in total? IF NONE, ENTER 0 -1 DON'T KNOW -2 REFUSED	|__|	
9.6d	How many ml is there in a **bottle** of STAINER chemical? -1 DON'T KNOW -2 REFUSED	|__|__|__|	
9.6e	How many **sachets** of STAINER chemicals did you use in total? IF NONE, ENTER 0 -1 DON'T KNOW -2 REFUSED	|__|__|	
9.6f	How many grams are there in a **sachet** of STAINER chemical? -1 DON'T KNOW -2 REFUSED	|__|__|__|	

Now I would finally like to ask you some questions still about education.

<table>
<tr><th>10</th><th colspan="3">EDUCATION</th></tr>
<tr><td>10.1a</td><td>What is the highest school grade completed by the farmer?</td><td>|__|__|</td><td>0 NONE
1 G1 2 G2 3 G3 4 G4
5 G5 6 G6 7 G7
8 G8 9 G9 10 G10 11 G11
12 G12 13 College
14 University 15 Other
-1 DON'T KNOW
-2 REFUSED</td></tr>
<tr><td>10.1b</td><td>What is the highest school grade completed by the first spouse?</td><td>|__|__|</td><td>99 NO SPOUSE
0 NONE
1 G1 2 G2 3 G3 4 G4
5 G5 6 G6 7 G7
8 G8 9 G9 10 G10 11 G11
12 G12 13 College
14 University 15 Other
-1 DON'T KNOW
-2 REFUSED</td></tr>
<tr><td>10.2a</td><td>How many BOYS of primary school age (7-13) are in this household? IF NONE, ENTER 0

-1 DON'T KNOW
-2 REFUSED</td><td>|__|__|</td><td>NUMBER OF BOYS 7-13 YEARS OLD
IF 0 GO TO 10.3a</td></tr>
<tr><td>10.2b</td><td>How many BOYS 7-13 years old in this household attend school at least 3 weeks per month when the school is in session? IF NONE, ENTER 0

-1 DON'T KNOW
-2 REFUSED</td><td>|__|__|</td><td>NUMBER OF BOYS 7-13 YEARS OLD</td></tr>
<tr><td>10.3a</td><td>How many GIRLS of primary school age (7-13) are in this household? IF NONE, ENTER 0

-1 DON'T KNOW
-2 REFUSED</td><td>|__|__|</td><td>NUMBER OF GIRLS 7-13 YEARS OLD
IF 0 GO TO SECTION 11</td></tr>
<tr><td>10.3b</td><td>How many GIRLS 7-13 years old in this household attend school at least 3 weeks per month when the school is in session? IF NONE, ENTER 0

-1 DON'T KNOW
-2 REFUSED</td><td>|__|__|</td><td>NUMBER OF GIRLS 7-13 YEARS OLD</td></tr>
</table>

The interview part is complete. Thank you for answering our questions.

11 COTTON PLOT MEASUREMENTS

I would now like to measure the LARGEST plot that was used to grow cotton during the 2012/2013 season.

RECORD GPS COORDINATES OF CLOSEST CORNER OF LARGEST COTTON FIELD

11.1	How many hectares (ha) is this plot? ENTER FARMER'S ESTIMATE (HA)	|__|__| . |__| HECTARE
11.1a	LATITUDE	S |__||__|º |__|__|.|__|__|__|
11.1b	LONGITUDE	E |__||__|__||º |__|__|.|__|__|__|
11.1c	ALTITUDE (M)	|__|__||__|__| . |__|
11.1d	MEASURED AREA OF 2012/2013 COTTON PLOT (HECTARES)	|__|__|.|__|__|__| HECTARES

YIELD ESTIMATION FOR THE LARGEST PLOT: COUNT THE NUMBER OF PLANTS AND BOLLS FOR 5 SAMPLING STATIONS WITH A LENGTH OF 3 METRES EACH, AND MEASURE THE DISTANCE IN CENTIMETRES TO THE ROW TO THE LEFT AND TO THE RIGHT

		NUMBER OF PLANTS	NUMBER OF BOLLS	DISTANCE (CM) TO ROW ON THE RIGHT	DISTANCE (CM) TO ROW ON THE LEFT
11.1e	SAMPLING STATION 1	|__|__|__|	|__|__|__|	|__|__|__|	|__|__|__|
11.1f	SAMPLING STATION 2	|__|__|__|	|__|__|__|	|__|__|__|	|__|__|__|
11.1g	SAMPLING STATION 3	|__|__|__|	|__|__|__|	|__|__|__|	|__|__|__|
11.1h	SAMPLING STATION 4	|__|__|__|	|__|__|__|	|__|__|__|	|__|__|__|
11.1i	SAMPLING STATION 5	|__|__|__|	|__|__|__|	|__|__|__|	|__|__|__|

I would now like to measure the SECOND LARGEST plot that was used to grow cotton during the 2012/2013 season.

RECORD GPS COORDINATES OF CLOSEST CORNER OF SECOND LARGEST COTTON FIELD

11.2	How many hectares (ha) is this plot? ENTER FARMER'S ESTIMATE (HA) -1 = DON'T KNOW -2 = REFUSED	|__|__| . |__|__| HECTARE
11.2a	LATITUDE	S |__||__|º |__|__|.|__|__|__|
11.2b	LONGITUDE	E |__||__|__||º |__|__|.|__|__|__|
11.2c	ALTITUDE (M)	|__|__||__|__| . |__|
11.2d	MEASURED AREA OF 2012/2013 COTTON PLOT (HECTARES)	|__|__|.|__|__|__| HECTARES

YIELD ESTIMATION FOR THE SECOND LARGEST PLOT: COUNT THE NUMBER OF PLANTS AND BOLLS FOR 5 SAMPLING STATIONS WITH A LENGTH OF 3 METRES EACH, AND MEASURE THE DISTANCE IN CENTIMETRES TO THE ROW TO THE LEFT AND TO THE RIGHT

		NUMBER OF PLANTS	NUMBER OF BOLLS	DISTANCE (CM) TO ROW ON THE RIGHT	DISTANCE (CM) TO ROW ON THE LEFT
11.2e	SAMPLING STATION 1	|__|__|__|	|__|__|__|	|__|__|__|	|__|__|__|
11.2f	SAMPLING STATION 2	|__|__|__|	|__|__|__|	|__|__|__|	|__|__|__|
11.2g	SAMPLING STATION 3	|__|__|__|	|__|__|__|	|__|__|__|	|__|__|__|
11.2h	SAMPLING STATION 4	|__|__|__|	|__|__|__|	|__|__|__|	|__|__|__|
11.2i	SAMPLING STATION 5	|__|__|__|	|__|__|__|	|__|__|__|	|__|__|__|

I would now like to measure the THIRD plot that was used to grow cotton during the 2012/2013 season.

RECORD GPS COORDINATES OF CLOSEST CORNER OF THIRD COTTON FIELD

11.3	How many hectares (ha) is this plot? ENTER FARMER'S ESTIMATE (HA)	|__|__| . |__|__| HECTARE
11.3a	LATITUDE	S |__||__|º |__|__|.|__|__|__|
11.3b	LONGITUDE	E |__||__|__||º |__|__|.|__|__|__|
11.3c	ALTITUDE (M)	|__|__||__|__| . I__|
11.3d	MEASURED AREA OF 2012/2013 COTTON PLOT (HECTARES)	|__|__|.|__|__|__| HECTARES

I would now like to measure the FOURTH plot that was used to grow cotton during the 2012/2013 season.

RECORD GPS COORDINATES OF CLOSEST CORNER OF FOURTH COTTON FIELD

11.4	How many hectares (ha) is this plot? ENTER FARMER'S ESTIMATE (HA)	|__|__| . |__|__| HECTARE
11.4a	LATITUDE	S |__||__|º |__|__|.|__|__|__|
11.4b	LONGITUDE	E |__||__|__||º |__|__|.|__|__|__|
11.4c	ALTITUDE (M)	|__|__||__|__| . I__|
11.4d	MEASURED AREA OF 2012/2013 COTTON PLOT (HECTARES)	|__|__|.|__|__|__| HECTARES

12	**INPUTS PROVIDED – TO BE COMPLETED BY COMPANY**	
12.1	HA CONTRACTED (COTTON)	**|__|**
12.2	SEED ISSUED (KG)	**|__|__|**
12.3	CHEMICALS ISSUED (Name x Quantity) – including foliar fertilizer and/or herbicides	**|______________________________| x |__|__|__| ml** **|______________________________| x |__|__|__| ml** **|______________________________| x |__|__|__| ml** **|______________________________| x |__|__|__| ml** **|______________________________| x |__|__|__| ml**
12.4	FERTILIZERS ISSUED	COMPOUND FERTILIZER NPK x 50 kg **|__|** NITROGEN TOP DRESSING x 50 kg **|__|**

13	**MARKETING – TO BE COMPLETED BY COMPANY**		
131	How many deliveries did the farmer make to the company?	|__|__|	
13.2	How much cotton did the farmers deliver to the cotton company?	|__|__|__|__|__|	KG
14	**BOLL WEIGHTS TO BE PROVIDED BY COMPANY**		
14.1	Average Boll weight for all PSUs		