Field Test Report on the Estimation of Crop Yields and Post-Harvest Losses in Ghana

Publication prepared in the framework of the Global Strategy to improve Agricultural and Rural Statistics

November 2017
Field Test Report on the Estimation of Crop Yields and Post-Harvest Losses in Ghana
# Table of Contents

Acronyms and Abbreviations ........................................................................ 4  
Acknowledgements ..................................................................................... 5  
Preface ........................................................................................................ 7  
1. Introduction and objectives .................................................................... 8  
2. Methodology and design ......................................................................... 10  
   2.1 Organization and responsibilities .................................................... 10  
   2.2 Scope of the survey ........................................................................... 10  
   2.3 Sampling design and selection ....................................................... 12  
   2.4 Data collection .................................................................................. 15  
   2.5 Objective measurements ................................................................. 16  
   2.6 The statistical estimation procedure .............................................. 24  
   2.7 Budget and resources ..................................................................... 28  
3. Data collection and field work ................................................................. 29  
   3.1 Training of the field teams ............................................................... 29  
   3.2 Deployment of the field teams ......................................................... 30  
   3.3 Monitoring and quality control ....................................................... 31  
   3.4 Challenges ...................................................................................... 32  
4. Data entry and compilation ..................................................................... 33  
   4.1 Data entry procedures and quality control ...................................... 33  
   4.2 Databases ......................................................................................... 34  
   4.3 Data processing, cleaning and imputation ...................................... 34  
5. Data analysis and main findings ............................................................... 36  
   5.1 Commodity coverage ...................................................................... 36  
   5.2 Farm sizes ........................................................................................ 36  
   5.3 Crop yields ....................................................................................... 37  
   5.4 Crop losses ...................................................................................... 40  
6. Conclusion ............................................................................................... 48  
7. Annexes .................................................................................................. 50  
   Annex 1: summary tables ....................................................................... 50  
   Annex 2: example of questionnaires ..................................................... 51  
   Annex 3: estimation of variances, standard deviations and confidence  
      intervals ............................................................................................. 52
# Acronyms and Abbreviations

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>APHLIS</td>
<td>African Postharvest Losses Information System</td>
</tr>
<tr>
<td>CSPro</td>
<td>Census and Survey Processing System</td>
</tr>
<tr>
<td>GAPS</td>
<td>Ghana Agriculture Production Survey</td>
</tr>
<tr>
<td>GS</td>
<td>Global Strategy to improve Agricultural and Rural Statistics</td>
</tr>
<tr>
<td>GSS</td>
<td>Government Statistical Service (Ghana)</td>
</tr>
<tr>
<td>HPHL</td>
<td>Harvest and Post-Harvest Losses</td>
</tr>
<tr>
<td>MoFA</td>
<td>Ministry of Food and Agriculture (Ghana)</td>
</tr>
<tr>
<td>PHL</td>
<td>Post-Harvest Losses</td>
</tr>
<tr>
<td>SARI</td>
<td>Savannah Agricultural Research Institute</td>
</tr>
<tr>
<td>SRID</td>
<td>Statistics, Research and Information Directorate (of MoFA)</td>
</tr>
</tbody>
</table>
Preface

The importance of post-harvest losses for food security has been appropriately reflected in the international development agenda. Within the Sustainable Development Goals (SDG) framework, Target 12.3 for Goal 12 strives to, “[b]y 2030, halve per capita global food waste at the retail and consumer levels and reduce food losses along production and supply chains, including post-harvest losses”. Similar objectives have also been adopted at regional level, with the example of the engagement taken by African heads of state in Malabo (June 2014) “to halve the current levels of post-harvest losses by the year 2025”1.

Given the strategic dimension of this topic and the lack of relevant reliable data and measurement methods, the improvement of methods for estimating post-harvest losses was identified by the member countries of the Food and Agriculture Organization of the United Nations (FAO) as a priority research topic. As such, it was included in the Research Component of the Global Strategy to improve Agriculture and Rural Statistics (GS). The objective of this research line is to develop cost-effective statistical methods for measuring post-harvest losses. To date, two methodological documents have been published: A literature review presenting the different measurement options2 and a gaps analysis.3 These documents are available on the GS website (http://www.gsars.org).

The measurement approaches and methods described in these publications needed to be tested to ensure their capability to produce quality results at a reasonable cost and their straightforward replication by countries. For this purpose, a pilot survey was conducted in Ghana from October 2016 to March 2017, testing a survey-based approach to measure harvest and post-harvest losses on the farm. This document describes the methodological approach that was tested and subsequently adopted. It also presents and discusses the data collected and the compiled indicators. Beyond the limitations and challenges inherent in any data collection exercise, the results presented in this report contribute to broadening the evidence base on food losses in sub-Saharan Africa. The methods used in this study are replicable by any country that wishes to improve its crop loss estimates. These methods will benefit from further testing,

1 Malabo Declaration on Accelerated Agricultural Growth and Transformation for Shared Prosperity and Improved Livelihoods. Malabo, Equatorial Guinea, 26–27 June 2014.
which is currently underway in other countries as part of the technical assistance activities of the GS.
Acknowledgements

This report has been prepared by Franck Cachia, international consultant for the Global Office of the Global Strategy (FAO). Mbaye Kebe, international consultant for the Global Office of the GS, provided key inputs and guidance for the design of the pilot survey implemented in Ghana. The Statistics, Research and Information Directorate of Ghana’s Ministry of Food and Agriculture (SRID/MoFA), which is in charge of the implementation of the survey, has contributed to several sections of this report. We are especially indebted to Mr Godsway Banini and James Ayittey, respectively Deputy Director and Head of the Data Management Team at SRID, for their active contribution to the preparation of this report and, more generally, for the successful implementation of the activities in Ghana.

We would also like to extend our thanks to Mr Harrison Opoku, Director of the SRID, for accepting to take part in this field test and for making his teams available. In this respect, we would like to express our gratitude to the SRID teams in Kintampo North and Sawla, the two districts chosen for this survey, and especially to the Directors, supervisors and enumerators without whom this field test could not have been undertaken.

Our final thanks go to the Savannah Agricultural Research Institute (SARI), and especially to Mr Issah Sugri, for supervising and conducting the laboratory analysis that were required in a successful and timely way.

Any errors or inconsistencies remaining in this document are the responsibility of the author.
Introduction and Objectives

The testing of measurement methods on crop losses at farm level is necessary to assess their relevance, cost-efficiency and replicability, before recommending them to countries.

The pilot survey undertaken in Ghana and described in this document concerns measurement of post-harvest losses (PHL)\(^4\) on the farm, a stage at which, in developing countries, losses are recognized to be high in relation to other stages of the production chain; in developed countries, losses are higher at the retail and consumption stages. Two different measurement methods were tested, assessed and compared: (1) farmers’ declarations and (2) objective measurements. Both approaches were conducted on a stratified sample of farmers selected at random. Therefore, the calculated indicators can be considered as representative of the targeted segment of farms. The use of a sample-survey approach enabled leveraging the farm surveys that are already conducted regularly in Ghana, in terms of the availability of listings, data and human resources.

Most sample surveys that have attempted to measure harvest and post-harvest losses (HPHL) at farm level rely on farmers’ declarations. Objective measurements through crop-cutting, weighting and laboratory analysis are not widespread, because of their comparatively high complexity and associated costs. However, the advantage of objective measurements, if done properly, is that they can yield more accurate results. This is true especially for complex topics such as losses, which farmers may find difficult to report, contrary to straightforward production surveys. Additionally, objective measurements can complement farmer declarations: they allow for crosschecking information and updating and calibrating technical parameters. They can also be used as a basis for modelling, by correlating measured losses with key drivers.

The measurement of yield and production is closely connected to the measurement of HPHL. Losses are often reported either as a percentage or as a quantity. When losses are expressed in percentage terms, they refer to harvested quantities; when expressed in quantitative terms, the harvested quantities must be used as the starting point from which losses can be discounted. Yield or

\(^4\) In this document, PHL comprises losses on the farm from (and including) harvest to (and including) storage.
production estimates are often based on farmers’ declarations, even though several studies have shown that these sources may be biased. In this context, providing quantitative evidence on the difference between farmers’ declarations and objective measurements regarding crop yields will aid towards understanding discrepancies, and possibly correcting them, while inevitably also contributing to improved loss estimates.

This report is articulated as follows: chapter 2 presents the survey and estimation methodologies; chapter 3 describes the organization of the field work; chapter 4 covers the data entry and cleaning work; chapter 5 discusses and compares the yield and loss indicators; and chapter 6 summarizes the main findings, their usefulness for future research and highlights the main challenges and limitations. Detailed tables and sample questionnaires are provided in the annexes.
Methodology and Design

2.1. Organization and responsibilities

The Global Office of the Global Strategy to improve Agricultural and Rural Statistics (GS) was in charge of the overall design, planning and supervision of the pilot survey. It also took the lead in establishing the estimation methodology and in carrying out the calculations from the farm-level data sets. Ghana’s Ministry of Food and Agriculture (MoFA) was in charge of data collection (through the deployment of supervisors and enumerators in two districts), data entry, quality control, and validation. This field test was also supported by the Savannah Agricultural Research Institute (SARI), which was in charge of the laboratory analysis, and by Ghana’s Statistical Service (GSS), which provided technical guidance and participated in the preparatory meetings and workshops.

2.2. Scope of the survey

Sector

The farms of the household sector were targeted to include essentially small- to medium-scale holdings producing both for own consumption and for the market. These farms were chosen because they are usually affected by high losses and because they tend to not keep complete records of their activity, leading to inaccurate reporting and data. Additionally, the data, information, sample frames and experience in survey-taking for the agricultural sector in Ghana, and at MoFA in particular, concerns the household sector.

Commodities

Four major cereal crops were chosen: maize, rice, millet and sorghum. These crops were available in the districts selected at the moment of the survey. The GS research on food losses currently focuses on cereals, although it will be expanded in 2018 to other commodities, including fruits and vegetables, milk, and animal proteins.
Losses

Losses were covered from harvest to storage on the farm. Losses for the following operations were measured through enquiry with the farmers: harvesting, threshing or shelling, cleaning or winnowing, drying, on-farm transport, and storage.

For objective measurements, the same steps were considered except for drying and transport operations. The practice followed by farmers in the chosen districts is to harvest the crops when they are already dry or close to dry, which means that losses during this stage should be minimal. The moisture content was recorded by the enumerator to ensure that comparisons can be made for a given moisture content. On-farm transport was also not considered because of the time and resources required to carry out these measurements. Furthermore, according to the experience of the field teams, the amounts lost at this stage are generally not significant for small farmers. Losses incurred during transportation from the farm to off-farm storage, processors, distributors or directly to the market may be more significant; however, they are beyond the scope of this study.

Agro-ecological zones and districts

Two districts in two different agro-ecological zones were chosen: Sawla-Tuna-Kalba district in the Northern Region, and Kintampo North, located in the transitional savannah. Sawla has only one cropping season for cereals (from June/July to October/November), while there are two seasons in the transitional savannah: a major season (from March/April to July/September) and a minor season (from August/September to November/December).
Timing

The fieldwork lasted from November 2016 to March 2017. The interviews and field measurements involving crop cutting were undertaken from November to January 2017. This period corresponded to the end of the harvesting season in the Northern Region and to the minor-season harvesting for maize and rice in the transitional savannah. The assessment of storage losses through objective measurements was conducted from January to March 2017.

2.3. Sampling design and selection

2.3.1. Overview

The primary objective of this pilot survey is to test the relevance and validity of the measurement approaches. While the sample size is too small to provide statistically representative results at the country or regional level, it is in principle large enough to draw reasonable inferences for the major crops in the two districts (cf. table 1).
The sample was selected on the basis of a random design, to limit the risk of large biases occurring in the results, according to a procedure familiar to the SRID/MoFA.

2.3.2. Sampling frame

The sample of farms was selected from the listings established for the 2015 Ghana Agriculture Production Survey (GAPS) in Kintampo North and Sawla-Tuna-Kalba. These listings contain the necessary information to characterize the holding according to the crops planted during the 2014–2015 agricultural season. The target population for this study comprises all farms in the two selected districts that planted at least one crop from among maize, rice, sorghum and millet during the 2015–2016 season.

2.3.3. Sampling selection and size

The farms were selected in two or three stages: in the first stage, 20 Enumeration Areas (EAs)\(^5\) were randomly selected from the 2015 GAPS listings in each district, with a probability of selection for each EA proportional to its size in terms of numbers of agricultural households. In the second stage, 14 holdings were randomly chosen from each selected EA using equal selection probabilities: all of them took part in the declarative survey. In a third stage, a subset of eight farms out of the 14 was selected using equal selection probabilities: at least one field in each of these eight farms was randomly selected for the objective measurements of yields and losses. Figure 2 illustrates the overall selection process.

\(^5\) The EAs correspond to the census blocks defined for the population census. They comprise a relatively small number of villages and households. In the two districts, the size of EAs varied considerably, from a minimum of 15 households to 208 households.
The listings in Kintampo and Sawla are relatively recent (having been created in 2015). However, even in such a limited time span, farmers may have exited the target population (disappearance of the household, change in crops cultivated, etc.) or new farmers may have entered. For this reason, preliminary verification work was required before the start of the survey. In practice, a sample of 14 farms in each EA was pre-selected from the 2015 listings and an additional 14 were kept as a reserve list. Enumerators visited the first set of 14 agricultural households to verify that they did grow at least one of the four crops covered by the study. If a household was found not to cultivate any of the four crops, or if the household no longer existed, it was replaced by a household from the reserve list. Table 1 provides additional details on the sample and its breakdown.

Table 1. Sample size, coverage and breakdown.
2.4. Data collection

2.4.1. Data collection mode

Paper questionnaires were used in this survey, and not tablets with embedded Computer-Assisted Programme Interview (CAPI) systems. This is because most of the interviewers, supervisors and survey coordinators had limited or no experience in using CAPI systems, and adding a further layer of complexity to an already complex survey was not advisable from a data quality point of view. In addition, the enumerators had to carry out several operations in the field (identify the fields, measure plot area, crop cutting, weighing, selecting and sorting samples for the laboratory, etc.) which in some cases are impractical to record on tablets.

2.4.2. Questionnaires

Nine questionnaires were used in this survey:

- FH0 – Identification form;
- FH1 – Identification of the subsample for the objective measurements;
- FH2 – Characterization of the agricultural household;
- FH3 – Production and losses: estimation by enquiry;
- FH4 – Field listing;
- FH5 – Field measurement;
- FH6 – Production and losses: estimation by objective measurements;
- FH7 – Storage losses: stock assessment; and
- FH8 – Storage losses: estimation by objective measurements.

Examples of filled-in questionnaires are provided in annex 2. Questionnaires FH5 to FH8 were administered only to the subsample of eight farmers in each selected EA.
2.5. Objective measurements

2.5.1. Field area measurement

For each of the eight farms selected in each EA for the objective measurements, a field with one of the four crops was randomly selected. The perimeter and area of this field was measured by the enumerators using a Global Positioning System (GPS). The enumerators were requested to mark and number each of the corners of the field, to report the length of each segment and to draw a sketch of the field on the questionnaire. The field teams were familiar with these operations. A proper measurement of the field area is necessary to extrapolate the yields and losses obtained from the subplot (see below, section 2.5.2) to the entire cultivated area for this crop, and then successively to the EA and district using the appropriate sample weights.

2.5.2. Placement of subplots

A subplot (or yielding plot, as it is known in Ghana), which serves as the reference for all of the objective measurements, was placed onto each selected field using the following procedure (see figure 3):

- Step 1: Identification of the field and of the starting point used for area measurement;
- Step 2: Marking and numbering (clockwise) of the corners of the field and measurement of the length of each segment;
- Step 3: Choice of a side or segment of the field through the selection of a random number between 1 and the number of corners of the field;
- Step 4: Choice of a point on the selected segment through the selection of a random number between 1 and the length of the segment;
- Step 5: Determination of the distance to be covered inside the field from the point of the segment, through selection of a random number between 1 and the half-perimeter of the field;
- Step 6: Enumerators’ entrance of the field perpendicularly from the point of the segment over the selected distance;
- Step 7: Placing of a subplot (6m x 6m square for maize, sorghum and millet, 3m x 3m for rice) at the point reached in step 6.
If this procedure leads to the placement of the subplot partly or entirely outside the bounds of the field, the enumerator repeats the procedure starting from step 3 until the subplot falls entirely inside the field. A random number table has been provided to the field teams to facilitate selection of the different random numbers.

**Figure 3. Placement of the yielding plot.**

### 2.5.3. Measurement of crop yields

The enumerators returned to the farm when the crop reached maturity to harvest the subplot, using the same practice as the farmer. This was to ensure that the results obtained would be as close as possible to the real yields and not to the potential or maximum yields. This principle remained valid for all the other measurements.

Once harvested, the produce of the subplot was bagged (according to local practices) and weighted. The yield was calculated by dividing these quantities by the area of the subplot (36 m\(^2\) for 6m x 6m subplots or 9 m\(^2\) for the 3m x 3m used for rice). This yield is used to estimate crop production for the entire field. For example, if 10 kg of maize have been harvested from the subplot and the field area is 1 ha, the estimated output of this field will be: 10 x 1 x \([10 000/36]\) = 2 778 kg.
2.5.4. Measurement of harvest losses

Immediately after completion of the harvesting, the enumerators returned to the subplot to collect and weigh the produce remaining on the ground (cobs, grains, ears, etc.). This amount corresponded to the quantities lost on the yielding plot at harvest. The quantities lost for the entire farm were estimated using the same procedure as for yields.

The produce remaining on the ground is usually present either in cobs or ears, or as loose grains. Adding the two together would lead to an overestimation of crop losses, as the empty cobs/ears have no or little economic and nutritional value beyond their use as animal feed. The enumerator is asked to report the cobs/ears and grain left on the ground separately. The cobs/ears are converted to a grain-equivalent amount using the ratio of grain to cobs/ears for each specific farm/field: it is obtained by dividing the grain obtained after threshing or shelling by the quantities of cobs/ears harvested. This conversion procedure has two limitations: the first is that it assumes that the cobs/ears remaining on the ground after harvest have the same grain weight as the harvested ones, when in fact it is likely that the lost or discarded ears/cobs have lost a significant amount of their grains. This will lead to an overestimation of the losses in grain-equivalent. The second limitation is that the grain weight obtained after threshing is, by construction, net of threshing losses. Using this amount to calculate the grain-to-cobs/ears ratio at harvesting is likely to lead to an underestimation of the quantities of grain lost at harvest. However, the first limitation (overestimation) is likely to dominate the second one (underestimation).

The quantity lost at harvest is obtained by summing up the weight of grain remaining on the ground after harvesting and the weight of cobs/ears in grain-equivalent. The percentage loss is calculated by dividing this amount by the sum of the quantities harvested and quantities lost at harvest.

2.5.5. Measurement of losses during post-harvest operations

Threshing or shelling

The harvest from the yielding plot is threshed or shelled (for maize) according to the method used by the farmer. After this process, the grain obtained and the discarded plant material (straw, etc.) are weighted separately. A sample of 250g of the discarded straw is taken and the grains in this sample are collected and weighted. This amount is then multiplied by the total weight of straw and divided by 250 to estimate the weight loss at threshing or shelling for the selected yielding plot.
Cleaning or winnowing

The grain obtained after threshing or shelling is then cleaned according to the method used by the farmer. After this process, the clean produce and the discarded unclean grain-straw mixture are weighted separately. A sample of 250g of the discarded grain-straw mixture is taken and the grains in this sample are collected and weighted. This amount is then multiplied by the total weight of the grain-straw mixture and divided by 250 to estimate the weight loss at cleaning/winning for the selected yielding plot. The measurement operations are illustrated in figure 4.

Several biases may affect the accuracy of these measurements. The most evident ones are described below:

- Post-harvest operations are not necessarily all performed immediately after harvest. In many cases, for example, the unthreshed or unshelled crop is stored in bundles, bags or left on the field (in heaps or stooks). The threshing or shelling is done whenever the household needs the produce to consume or to sell. In this experiment, the harvest and post-harvest operations were done in sequence; therefore, the loss estimates may not be fully representative of the actual farming practices followed in the two districts;

- Some farmers, especially in the district of Kintampo North, commonly use machines to shell, clean and dehusk maize. In this experiment, it was
assumed that farmers carried out these operations manually. Therefore, the results are not fully representative of the farming practices commonly used in Kintampo North district.

2.5.6. Losses during storage

The stored quantities were estimated in different ways: all farmers of the sample were asked to report the quantities stored from the current or past harvest; for the farmers selected for physical measurements, the stocks for the selected crop were estimated by asking the farmer to report the amount in store at the time of each visit and any additions or withdrawals between each visit. Three visits were made, once every month, during which samples of grain were taken and sent to SARI’s laboratory for analysis. Farmers were not compensated, neither financially nor in kind, for the samples of grain taken from their stock.

Measurements and laboratory analysis

Immediately after the harvesting of the yielding plot and the completion of the various post-harvest operations, the enumerators were asked to take a first sample of grain from the storage facility. The teams followed a specific procedure: if possible, they selected two samples (1kg–2kg) from two different observational units (generally a bag) using multicompartment spears. When simple spears were used, enumerators were advised to pick samples at different points of the bag. Then, the two samples had to be homogenously merged, for example through coning and quartering, and a new sample (1kg–2kg) was selected from this homogeneous mixture. This sample was then sent to SARI’s laboratory for analysis. This procedure was repeated two additional times at monthly intervals: each storage facility in the selected farm was therefore visited three times over a three-month period and three samples of the same crop were sent for laboratory analysis.

This procedure is relatively lengthy because it must to cover a sufficiently long storage period to capture meaningful losses. The time period should be chosen in function of the storage practices prevailing in the selected country; however, three months is generally considered as the minimum. A period of six to nine months covering the entire agricultural year would have been preferable; however, this was not possible due to time and budget constraints. Experienced enumerators should be assigned to this task, as taking grain samples in the required manner can be complex; the fieldwork should be well planned and organized so that the samples can reach the laboratory as quickly as possible. The activity is therefore complex and relatively costly; however, it does not have to be carried out each year as loss percentages during storage are unlikely to vary significantly from year to year under normal conditions. As an indication, the
costs related to this activity represented approximately 30 percent of the total fieldwork expenses of this field test.

Figure 5. Weighing of a sample of millet (SARI, February 2017).

The following measurements were carried out:

- Weighing of the grain samples received;
- Measurement of moisture content;
- Separation of damaged from undamaged grains;
- Identification of the cause of damage (insect, fungi, other);
- Counting and weighing damaged and undamaged grain.

The results of these analyses have been recorded in the forms provided for this purpose. A detailed guide for laboratory assistants has been prepared and provided to SARI to ensure that they will be undertaken as indicated and in a consistent way.
Note: These measurements and the resulting loss estimates must be interpreted with care. Some of the limitations affecting these measurements are discussed below:

- A perfectly random sample of grains is not easy to collect. First, the produce may not be stored in bags but in bundles, as is often the case for millet and sorghum in traditional farms, or in large granaries or silos. This complicates the selection of a random sample. Second, the grain that is stored may come from the previous harvest (especially in Kintampo North, that has two seasons) or may not come from the farmer’s field. The stored units (bags, etc.) must be properly identified before selecting samples.

- Grain samples need to be sent to the laboratory for analysis. This means that if the produce is stored in cobs/ears and threshed by the household on an as-needed basis, the enumerator will have to thresh the crop him- or herself before picking the sample and sending it to the laboratory. The characteristics of the grain (moisture content, pest infestation, etc.) may be artificially altered, compared to what they are in practice when stored unthreshed or unshelled. The resulting loss estimates may therefore not be fully representative of the actual post-harvest and storage practices. This potential bias could have been partially mitigated by stratifying the farms according to storage practices (threshed/unthreshed, for example) and estimating losses separately for each stratum.

- The physical characteristics of the produce might have been altered from the time the sample was taken from the storage facility to the moment it reached the laboratory. The teams were instructed to dispatch the samples as quickly as possible and to the extent possible, given the
logistical arrangements in place and the distance between the laboratory and the two districts.

**Calculation of percentage losses**

Percentage losses during storage are directly calculated, using the laboratory measurements. The count and weight method is used, based on the formula proposed by Harris and Lindblad (1978):

$$ l_s^{(t)} = \frac{1}{W_u} \left[ \frac{N_d}{N} W_u - \frac{N_u}{N} W_d \right] $$

where:

- $l_s^{(t)}$ is the percentage loss estimated for a given household visited in month $t$;
- $N_u$ is the number of undamaged grains ($W_u$ the corresponding weight);
- $N_d$ is the number of damaged grains ($W_d$ the corresponding weight); and
- $N = N_u + N_d$ is the total number of grains in the sample.

A more intuitive version of this formula can be determined by using the proportionality between the weight of each portion of the grain sample (damaged and undamaged) and its size in terms of number of grains: $W_u = \alpha_u N_u$ and $W_d = \alpha_d N_d$, with $\alpha_u$ (respectively $\alpha_d$) the average weight of an undamaged grain (resp. damaged). The following inequalities should hold: $\alpha_u > 0$, $\alpha_d > 0$ and $\alpha_u > \alpha_d$ (on average, an undamaged grain should weigh more than a damaged grain). Using these notations, it is possible to show that:

$$ l_s^{(t)} = \frac{N_d}{N} \left[ \frac{\alpha_u - \alpha_d}{\alpha_u} \right] $$

The percentage storage loss is equal to the percentage difference between the average weights of undamaged and damaged grain, weighted by the share of damaged grains in the total number of grains of the sample. For example, if damaged grains weigh on average 25 percent less than undamaged grains and if damaged grains represent 50 percent of the sample, the percentage loss estimated for this sample will be 50 percent $\times$ 25 percent $= 12.5$ percent. This formula shows that this method accurately measures weight losses, but disregards qualitative losses.
2.6. The statistical estimation procedure

The primary objective of this study is to measure percentage (or relative) losses, and not absolute losses expressed in physical units. Percentages are generally more stable and less prone to biases than absolute indicators. In any case, absolute losses can be obtained from percentages by applying the loss percentages to the harvested quantities, provided that the latter are properly measured. This section presents this study’s main survey variables and parameters of interest and describes the statistical procedure used for their estimation.

2.6.1. Parameters of interest and survey variables

The objective of this study is to estimate average percentage losses for:

- Harvest

- Post-harvest operations, broken down by type of operation:
  - Threshing or shelling
  - Cleaning or winnowing
  - Drying (only for the declarative survey)
  - Transport (only for the declarative survey)
  - On-farm storage

- Two aggregates: (i) PHL; and (ii) HPHL (known as post-production losses)

These parameters are estimated by crop and by type of measurement method (objective measurements and farmers’ declarations), except for transport and drying losses, which are only based on farmers’ declarations. The survey variables used to estimate the parameters of interest are described in table 2.
Table 2. Variables, notations and formulae.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Absolute (kg)</th>
<th>Relative (%)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Quantities:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harvested</td>
<td>(H)</td>
<td></td>
<td>All variables refer to a single agricultural household / holding</td>
</tr>
<tr>
<td><strong>Brought to:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Threshing/shelling</td>
<td>(T)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cleaning/winnowing</td>
<td>(C)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drying</td>
<td>(D)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transportation</td>
<td>(T)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Storage</td>
<td>(S)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Losses during:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harvesting</td>
<td>(L_H)</td>
<td>(l_H = \frac{L_H}{H + L_H})</td>
<td>(H + L_H) is a measure of potential harvested quantities</td>
</tr>
<tr>
<td>Threshing/shelling</td>
<td>(L_T)</td>
<td>(l_T = \frac{L_T}{T})</td>
<td></td>
</tr>
<tr>
<td>Cleaning/winnowing</td>
<td>(L_C)</td>
<td>(l_C = \frac{L_C}{C})</td>
<td></td>
</tr>
<tr>
<td>Drying</td>
<td>(L_D)</td>
<td>(l_D = \frac{L_D}{D})</td>
<td></td>
</tr>
<tr>
<td>Transport</td>
<td>(L_{Tr})</td>
<td>(l_{Tr} = \frac{L_{Tr}}{T_{Tr}})</td>
<td></td>
</tr>
<tr>
<td>Storage (farmers’ declarations)</td>
<td>(L_S)</td>
<td>(l_S = \frac{L_S}{S})</td>
<td></td>
</tr>
<tr>
<td>Storage (objective measurements)</td>
<td></td>
<td>(l_S^{(t)})</td>
<td>(t = 3) visits; (l_S^{(t)}) is the percentage storage loss at visit (t) calculated using the count and weight method.</td>
</tr>
<tr>
<td><strong>Aggregates:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHL</td>
<td>(L_{PH} = L_T + L_C + L_S)</td>
<td>(l_{PH} = \frac{L_{PH}}{H})</td>
<td></td>
</tr>
<tr>
<td>HPHL</td>
<td>(L_{HPH})</td>
<td>(l_{HPH} = \frac{L_{PH} + L_H}{H + L_H})</td>
<td></td>
</tr>
</tbody>
</table>
2.6.2. Estimation procedure

Average crop yields

For a given crop and district, the average yield $\hat{Y}$ (in kg/ha) corresponds to the ratio of the estimated harvest to the estimated planted area for this district. This can be calculated by calculating the average of the estimated yields for each EA weighted by the estimated share of each EA in the district’s crop area:

$$\hat{Y} = \sum_j \theta_j(ha) \cdot \hat{Y}_j$$

where:

$$\hat{Y}_j = \frac{\hat{H}_j}{\hat{HA}_j} \text{ and } \theta_j(ha) = \frac{HA_j}{HA}$$

with:

$$\hat{H}_j = \sum_{i \in S_j} w_i \cdot H_i \text{ and } \hat{HA}_j = \sum_{i \in S_j} w_i \cdot HA_i$$

The variable $w_i$ is the sample weight attached to household $i$ (see box 1 for details on how these weights were obtained), $HA_i$ is the area planted by this household for the selected crop and $S_j$ is the sample of households in EA$j$.

The average yield for the two districts is calculated by taking the average of the two yield estimates, weighted by the respective share of each district in the total estimated crop area.

Average crop losses

The estimation procedure is the same for percentage crop losses ($\hat{l}$):

$$\hat{l} = \sum_j \theta_j(q) \cdot \hat{l}_j$$

where $q$ is the denominator relevant for the type of loss measured. For example: $q = H + L_H$ for harvest losses, $q = T$ for threshing losses and so on for the other
operations (see table 2 above for the full list). Further details on the estimation procedure are available in Jha *et al.* (2015), among others.

**Variances, standard deviations and confidence intervals**

Confidence intervals have been calculated for the key parameters. These reflect the variance in the average estimates due to the sample selection process. The details of the calculation are provided in annex 1.

**Box 1. Determination of sample weights.**

The sample weight of a given household is equal to the inverse of the probability of selecting this household in the sample: \( w_i = 1/\pi_i \). The weights are different for the estimates based on objective measurements because these estimates are based on a subsample of the households in each EA. The selection probabilities are the following:

Households selected for inquiry only:

\[
\pi^{(1)} = \frac{20 \frac{N_j}{N}}{\text{probability of selecting a farm in EA} \ j} \cdot \frac{14}{N_j} \]

Households also selected for objective measurements:

\[
\pi^{(2)} = \pi^{(1)} \cdot \frac{8/14}{\text{probability of selecting a farm for the objective measurements in EA} \ j} = \pi^{(1)} \cdot \frac{4}{7}
\]

with \( N_j = \sum_{i \in S_j} N_i \) being the total number of farms in EA \( j \) and \( N = \sum_j N_j \) being the total number of farms in the district.

Had the number of households in each EA remained as initially planned (14), the weights would have remained unchanged and equal in each district. However, due to partial nonresponse and to the rejection of certain records during the data analysis stage, the number of usable households in certain EAs is sometimes lower than initially targeted. The initial weights have therefore been adjusted as follows: \( \bar{w}_j = \frac{14}{n_j} \, w \), where \( n_j \leq 14 \) is the effective number of usable households in each EA.

---

2.7. Budget and resources

The budget allocated to this field test was USD 84 945. This amount covered all costs related to collection, including the training of enumerators and data entry (see table 3 below); however, it excluded the time worked by GS consultants and the cost of their missions to Ghana. The survey costs per sample unit surveyed amounts to USD 152, which makes it a relatively costly survey for a developing country such as Ghana. This is due to the complex and lengthy field experiments that involve crop-cutting and other farm operations. If such a survey were to be implemented at full scale, the unit cost would be lower because some of the cost items – such as training, data entry and processing – are not perfectly proportional to the size of the sample.

<table>
<thead>
<tr>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training of data collection teams</td>
</tr>
<tr>
<td>Fieldwork</td>
</tr>
<tr>
<td>Data entry, validation and processing</td>
</tr>
<tr>
<td>Laboratory analysis</td>
</tr>
<tr>
<td><strong>Total</strong></td>
</tr>
</tbody>
</table>

Note: All amounts in USD.

<sup>7</sup> Includes daily subsistence allowances for each participant, trainer and accompanying staff (drivers, administrative staff), fuel costs, hiring of training facilities, stationery, etc.

<sup>8</sup> Includes daily subsistence allowances for enumerators, supervisors and drivers, as well as hiring of vehicles, fuel costs, insurance coverage, printing of questionnaires and miscellaneous expenses.
Data Collection and Field Work

3.1. Training of the field teams

Two trainings were organized for the data collection teams – one in Kintampo North and one in Sawla\(^9\) in October 2016. The participants were the field staff (approximately ten enumerators and two supervisors per district), the district-level management teams of MoFA and key staff from the SRID headquarters. The trainings were led and facilitated by GS consultants and by Mr Godsway Banini, Deputy Director of SRID.

During the trainings, the nine questionnaires were each presented in detail and discussed with the participants. In addition, field measurement techniques were presented, discussed and demonstrated, such as the technique for field measurement and placement of the yielding plot. The questionnaires were also pilot-tested on two farmers during interviews lasting one to two hours each.

Although the teams consisted mostly of experienced enumerators and supervisors, there was no significant experience in PHL surveys. The trainings were therefore useful in clarifying some of the most complex concepts, as well as in providing additional insights on how to conduct the interviews and field measurements to reduce potential biases. Methods that were new to the teams were also presented (such as the approach to placing a yielding plot on a field) and were generally well understood and adopted.

The participants provided relevant feedback to the GS consultants regarding the content and structure of the questionnaires, especially its relevance to the context of Ghana: measurement units (bags,\(^10\) etc.) were adapted to reflect local market conditions and some modules were restructured. For example, the existence of two cropping seasons in Kintampo justified the inclusion of an additional item.

\(^9\) For practical reasons, the training for the data collection teams that covered Sawla was organized in Bole, the neighbouring district.

\(^10\) The field teams were familiar with the kg equivalents of the nonstandard units. According to the teams, such units adequately reflected actual farming practices.
in the questionnaire to specify whether the data reported by the farmer referred to the current or the past season.

3.2. Deployment of the field teams

This section describes how the fieldwork was organized to collect the required data in the selected districts and EAs.

In each district, two teams made up of five enumerators and one supervisor were formed. Each team collected data from ten EAs. A total of 20 EAs were therefore covered in each district. The overall field exercise was coordinated by the municipal or district Directors of Agriculture for each district. Hence, the data collection team in each district consisted of thirteen officers from MoFA.

The field data collection activities in the districts were preceded by a sensitization exercise. The objective was to alert the farmers and the wider community to the PHL survey and solicit their support.

The entire exercise was divided into four phases, as illustrated in table 4 below, to ensure proper planning of the data collection activities. During each phase, the supervisors visited the field at least three times.
Table 4. Work plan for data collection and measurement activities.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Act. No.</th>
<th>Activity</th>
<th>Start date</th>
<th>End date</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>Sensitization/verification of selected holders</td>
<td>7 November 2016</td>
<td>12 November 2016</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>Household enquiry</td>
<td>15 November 2016</td>
<td>26 November 2016</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>Field area measurements</td>
<td>28 November 2016</td>
<td>31 December 2016</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>Placement of the yield plots</td>
<td>28 November 2016</td>
<td>28 January 2017</td>
</tr>
<tr>
<td>3</td>
<td>A</td>
<td>Harvesting of the yield plots</td>
<td>28 November 2016</td>
<td>28 January 2017</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>Determination of losses from the yield plots</td>
<td>28 November 2016</td>
<td>28 January 2017</td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>Sampling of harvested stored grains for laboratory analysis</td>
<td>January 2017</td>
<td>March 2017</td>
</tr>
</tbody>
</table>

3.3. Monitoring and quality control

During the fieldwork, SRID undertook two monitoring visits to the two districts. Two teams made up of four members of staff each were allocated to each of the two districts. FAO-Ghana also participated in these visits to monitor the field operations in the two districts.

The teams’ terms of reference for the monitoring visits included:

- Checking completed questionnaires;
- Choosing randomly an EA to visit; and
- Discussing with field enumerators the challenges they face and provide solutions, where possible.

The first visit was in November 2016. The monitoring team edited completed questionnaires (FH0, FH1, and FH2). The team also discussed all necessary corrections with the field enumerators and supervisors and advised them to take
some of the questionnaires back to the field for re-filling. The team then took the completed questionnaires without errors to Accra for data entry.

The second monitoring was in January 2017. The monitoring team edited the completed questionnaires (FH3, FH4, FH5 and FH6). All anomalies detected were notified to the field enumerators. The team randomly chose an EA and crosschecked the yield plots established.

The survey was also closely supervised by district Directors and Supervisors. They checked the completed questionnaires during field visits, observed the enumerators during questionnaire administration and provided technical backstopping where necessary.

At the end of each phase of data collection, enumerators submitted their forms to the supervisors for crosschecking and data quality control before submission to the SRID headquarters in Accra. In some cases, the data were cleaned and enumerators were asked to return to the field to repeat some data collection and measurement activities. During one of the supervisors’ field visits, it was detected that the weighing scale provided was unable to weigh the 100 grains of millet. A more sensitive scale was found to continue with the work and ensure data quality.

### 3.4. Challenges

The execution of the PHL survey had to face a number of challenges, the major ones being the following:

- **The late start of the survey:** in Sawla, part of the maize fields had been harvested before the start of the survey. Additional fields had to be selected to make up for these missed fields.

- **The weighing scale was not sensitive enough to weigh chaff and fallen grains,** such as a quantity of 100 millet grains. This affected the quality of some of the measurements.

- **Issues with some sections or questions of forms FH3 and FH7 did not conform with the realities on the field,** and therefore needed to be corrected and adapted.

- **Sociocultural practices such as funerals and other rural engagements negatively affected planned data collection and measurement activities.**
Data Entry and Compilation

4.1. Data entry procedures and quality control

The Census and Survey Processing System (CSPro), a standard software developed and maintained by the U.S. Census Bureau, was used for data entry.

The data, collected on paper forms, was entered into CSPro on a continuous basis, as the clean and validated questionnaires were received from the two districts. The data entry started with the enumeration forms (FH0 and FH1) and ended with the forms on the measurement of storage losses through enquiries (FH7) and laboratory analysis (FH8). This process enabled reduction of data entry time compared to an approach in which the information contained in all the questionnaires had to be entered at once.

The data entered into CSPro underwent a range of validation checks, most of which were directly implemented in the data entry masks. These checks consisted first in ensuring that the identifiers (at plot, field, household, EA and district levels) had been correctly entered. In addition, several consistency rules were applied to identify issues in the reporting of the information by the enumerators. For example, the total farm area reported was compared to the sum of the area reported for the four crops. Additional rules checked whether the reported quantities harvested were larger than the quantities handled at the different processing stages (threshing, drying, etc.).

The application of these rules led to the identification of a certain amount of inconsistencies and the rejection of some questionnaires. These questionnaires were sent back to the supervisors for clarification. In some cases, the supervisors were asked to collect the required information from the farmers again.

Although this process delayed data entry and the final delivery of the data sets, it ensured a minimum level of data quality. The effort devoted to this work might have been considerably reduced if the data had been collected with an appropriate CAPI application, with built-in consistency and validation checks. However, CAPI was not used in this survey for reasons described in section 0.
4.2. Databases

After completion of the data entry stage, databases in .csv and .xlsx formats were generated and sent to the GS. Each form (FH0 to FH8) is associated with three data sets: the first containing basic information on the survey (name of surveyor, identity number, date of survey), a second with the data collected from farmers and a third containing the metadata for each variable (variable code, description, etc.). As the enumerators visited the same household three times for the storage loss assessments, three sets of data are available for forms FH7 and FH8 – one for each visit.

The data was collected for four different statistical units: households (forms FH0, FH1, FH3, FH7 and FH8), household members (FH2), fields (FH4 and FH5) and yielding plots (FH6). As identifiers have been given to each unit, relationships can be established between each of them. For example, losses obtained from the yielding plot (FH6) can be extrapolated to the measured field area (FH5) and to the entire cultivated area for this crop in the concerned household (FH4), using plot, field and household identifiers as matching variables.

4.3. Data processing, cleaning and imputation

Despite the care that was exercised in data collection and the validation and consistency checks applied during data entry and the field verifications, a certain amount of inconsistent data, missing values or incorrect information remained to the final data sets. To mitigate their impact on the quality of the final indicators, the data was cleaned and, when possible, the missing information was imputed.

These operations included:

- Imputation of the missing data on the weight in kg of nonstandard units (cocoa bags, koko bowl, etc.) using the median unit weight;
- In form FH6 (loss assessment through measurements), part of the data was reported in kg when grams were expected: these cases were identified and the incorrect data converted back to grams;
- Consistency rules were applied to data from form FH6: the harvested quantities should be higher than the losses at harvest, the weight of grains in 250g of straw after threshing cannot be higher than 250g, etc. The 24
households that did not pass these tests were eliminated from the final data set;

• The largest outliers (top and bottom 2.5 percent of the observations) for some of the key variables (area planted, losses measured at harvest, threshing, etc.) were identified and eliminated;

• The sample weights were adjusted to account for reductions in effective sample size due to the rejection of some records and to partial nonresponse. The adjustment method used was simple calibration, adjusting weights by the ratio of targeted sample size to the sample size of the responding farms, as described in box 1 above on the determination of sample weights.
Data Analysis and Main Findings

5.1. Commodity coverage

A total of 560 farms were visited and interviewed, equally distributed across the two districts. Of the 560 farms interviewed, 475 had planted maize (table 5), the country’s major crop. Sorghum and rice fields were found in 287 and 135 farms respectively, mostly located in the district of Sawla. Millet was present in 190 farms, almost exclusively in Sawla.

Table 5. Distribution of the farm sample by planted crops.

<table>
<thead>
<tr>
<th>Crops</th>
<th>All districts</th>
<th>Kintampo</th>
<th>Sawla</th>
</tr>
</thead>
<tbody>
<tr>
<td>Millet</td>
<td>190</td>
<td>2</td>
<td>188</td>
</tr>
<tr>
<td>Maize</td>
<td>475</td>
<td>249</td>
<td>226</td>
</tr>
<tr>
<td>Rice</td>
<td>135</td>
<td>44</td>
<td>89</td>
</tr>
<tr>
<td>Sorghum</td>
<td>287</td>
<td>47</td>
<td>240</td>
</tr>
</tbody>
</table>

5.2. Farm sizes

The median planted area by farm varied from 0.4 to 1.2 ha, depending on the district and crop (table 6). The median cultivated areas were lowest for rice (0.4 ha) and highest for sorghum (1.2 ha). The planted areas in Sawla tend to be larger than in Kintampo (figure 7).
5.3. Crop yields: farmers’ declarations

Farmers provided own estimates of quantities harvested (in nonstandard units) and area planted (in acres) for their main crop. The quantities were converted to kg using conversion factors provided by the SIRD. Crop yields in kg/ha were estimated for each district according to the procedure described in section 0 above. Results for rice in both districts and sorghum and millet in Kintampo should be interpreted with care given the small sample sizes (for the latter, given that only two farms were concerned, results were not disseminated).

On average, over the two districts, rice yields – estimated at 1.5 tons per hectare (t/ha) – are superior to those of other cereals (table 7). The yields for maize and sorghum are estimated at just under 1 t/ha. For sorghum, the average yields in

Table 6. Median cultivated areas by farm.

<table>
<thead>
<tr>
<th>Hectares (ha)</th>
<th>All districts</th>
<th>Kintampo</th>
<th>Sawla</th>
</tr>
</thead>
<tbody>
<tr>
<td>Millet</td>
<td>0.8</td>
<td>-</td>
<td>0.8</td>
</tr>
<tr>
<td>Maize</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>Rice</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>Sorghum</td>
<td>1.2</td>
<td>0.4</td>
<td>1.2</td>
</tr>
</tbody>
</table>
Kintampo are significantly higher than in Sawla. Overall, these yields are consistent with recent estimates calculated by the SIRD through the GAPS.

Taking the average yield estimated for the two districts, the confidence bands indicate a limited dispersion of the estimates around the average, except for rice. For this crop, which is found in relatively few of the sampled households, there is an estimated 90 percent chance that the yield is between 1.4 and 2.7 t/ha.

Table 7. Crop yields from farmers’ declarations.

<table>
<thead>
<tr>
<th>Main crop</th>
<th>Averages</th>
<th>Confidence intervals (90%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All districts</td>
<td>Kintampo</td>
</tr>
<tr>
<td>Millet</td>
<td>562</td>
<td>-</td>
</tr>
<tr>
<td>Maize</td>
<td>969</td>
<td>984</td>
</tr>
<tr>
<td>Rice</td>
<td>1,472</td>
<td>1,513</td>
</tr>
<tr>
<td>Sorghum</td>
<td>833</td>
<td>1,584</td>
</tr>
</tbody>
</table>

5.3.2. Objective measurements

The procedure adopted by enumerators to measure yields from the selected fields is discussed in section 2.5.3 above. The sample of households selected for yield and loss measurements is a subsample: in each EA, eight out of the 14 households were selected for objective measurements. In addition, as discussed in section 4.3, additional records were excluded due to insufficient data quality. The number of available records to calculate crop yields was therefore considerably lower than for declaration-based estimates, especially for millet, rice and sorghum (table 8). The results presented in this section should therefore be interpreted with care.

Table 8. Number of fields by crop.

<table>
<thead>
<tr>
<th>Crops</th>
<th>Number of fields by crop</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All districts</td>
</tr>
<tr>
<td>Millet</td>
<td>20</td>
</tr>
<tr>
<td>Maize</td>
<td>126</td>
</tr>
<tr>
<td>Rice</td>
<td>13</td>
</tr>
<tr>
<td>Sorghum</td>
<td>52</td>
</tr>
</tbody>
</table>
Table 9. Crop yields from objective measurements.

<table>
<thead>
<tr>
<th>Crops</th>
<th>Averages</th>
<th>Confidence intervals (90%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All districts</td>
<td>Kintampo</td>
</tr>
<tr>
<td>Millet</td>
<td>1 068</td>
<td>-</td>
</tr>
<tr>
<td>Maize</td>
<td>2 272</td>
<td>2 270</td>
</tr>
<tr>
<td>Rice</td>
<td>1 887</td>
<td>2 182</td>
</tr>
<tr>
<td>Sorghum</td>
<td>1 430</td>
<td>1 065</td>
</tr>
</tbody>
</table>

5.3.3. Farmers’ declarations versus objective measurements

The average crop yields estimated through objective measurements are larger than those reported by farmers. The difference is marked for maize, with 2.3 t/ha compared to 1.0 t/ha, respectively (table 10). The two different methods provide a similar ranking of the crops according to their yield: maize and rice have the highest yield, sorghum and millet the lowest. Objective measurements place maize as the highest yielding crop, and not rice as for the estimates based on farmer reporting.

As illustrated by the confidence bands, the precision is lower for objective measurements. This is because of the smaller sample size and, probably, also because of the higher complexity of the measurement operations. Indeed, the latter increase the risk of mistakes and differences among field teams in the way operations are conducted, leading to a higher dispersion of the results. The overlapping confidence bands for rice indicate that the difference in yields between the two methods is not statistically significant, contrary to the other crops.

The fact that yields estimated from objective measurements tend to be higher than those based on farmer reporting is in line with the findings of the recent literature, several studies having reached similar conclusions for different crops, agroclimatic zones and study designs. Verma et al. (1998)\(^\text{11}\) provides comprehensive evidence on yield estimation for five African countries (Benin, Central African Republic, Kenya, Niger and Zimbabwe). One of their findings is that production estimates based on crop-cutting techniques are on average 30 percent higher than farmer-reported production.

Some studies have investigated the causes of this difference. The results of a recent research project in the United Republic of Tanzania, led by the GS,

indicated that one of the possible explanations is that production estimates based on farmers’ declarations are relatively accurate while plot areas are not: farmers tend to overestimate the size of their plots, resulting in underestimated yields.\footnote{More details on this study are available by contacting the authors of this report.}

\begin{table}[h]
\centering
\caption{Crop yields (average for the two districts): Farmers’ declarations versus objective measurements.}
\begin{tabular}{|l|c|c|}
\hline
          & Farmers’ declarations     & Objective measurements     \\
\hline
Millet    & 562 [559 – 621]           & 1 068 [959 – 1 275]       \\
Rice      & 1 472 [1 402 – 2 682]     & 1 887 [1 190 – 3 212]     \\
Sorghum   & 833 [825 – 974]           & 1 430 [1 308 – 1 662]     \\
\hline
\end{tabular}
\end{table}

5.4. Crop losses

5.4.1. Farmers’ declarations

Percentage losses for the various harvest and post-harvest operations were estimated on the basis of farmers’ declarations for their main crops. The estimation procedure is explained in section 0 above.

Estimates of HPHL over the two districts (table 11) range from 2.9 percent (sorghum) to 9.5 percent (maize). The confidence intervals indicate a relatively low dispersion of the results, with the exception of rice, as already noted in the case of yields. The larger standard deviations obtained for rice are due partly to lower sample sizes and probably also to differences in farming practices and technologies. Additional investigations are necessary to test this hypothesis.

There is a significant difference between the two districts, with higher losses occurring in the district of Kintampo for all four crops of this study. This can be explained, at least partially, by the differences in climate conditions: in Kintampo, where rainfall is higher than in Sawla, crops generally require additional drying and are more likely to suffer from degradation and losses due to a higher moisture content. Conversely, crops in Sawla are generally harvested dry and quickly moved to the farm’s storage facility. The magnitude of the gap between the two districts may also suggest a potential lack of uniformity in how the data on losses was obtained from the farmers.
Table 11. Harvest and post-harvest losses (farmers' declarations).

<table>
<thead>
<tr>
<th>Main crop</th>
<th>Averages</th>
<th>Confidence intervals (90%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All districts</td>
<td>Kintampo</td>
</tr>
<tr>
<td>Millet</td>
<td>5.0</td>
<td>12.4</td>
</tr>
<tr>
<td>Maize</td>
<td>9.5</td>
<td>12.3</td>
</tr>
<tr>
<td>Rice</td>
<td>5.1</td>
<td>5.2</td>
</tr>
<tr>
<td>Sorghum</td>
<td>2.9</td>
<td>11.1</td>
</tr>
</tbody>
</table>

The breakdown by task shows that losses occurring during harvest are highest (figure 8), immediately followed by threshing/shelling losses, in a measure far greater than those sustained during transport and storage. It is worth noting that given that the quantities handled during harvesting are greater than or equal to the quantities handled at later stages (threshing, drying, etc.), this result also holds for losses in quantity terms. This predominance of harvest losses over other types of losses is confirmed by the African Postharvest Losses Information System (APHLIS)\(^\text{13}\), which provides typical percentage harvest losses of 4–8 percent, above other links of the post-harvest chain.

Rice stands out because its losses are the highest during cleaning. This can be explained by the specificity of the post-production process for this crop, which involves several operations before becoming ready for consumption or selling. Once threshed, the paddy may undergo several stages of processing, such as winnowing, de-stoning and milling (dehusking, polishing, etc.), with losses occurring at each stage. It is possible that these losses have all been included under the “Cleaning” heading, as there was no other possible choice in the questionnaire.

\(^{13}\) [www.aphlis.net](http://www.aphlis.net)
5.4.2. Objective measurements

The loss estimates based on objective measurements are provided as averages for the two districts (table 12). The estimation procedure is detailed in section 0 above.

Average harvest losses for maize are estimated at 9.0 percent, followed by threshing (3.4 percent), cleaning (4.5 percent) and storage (4.1 percent). For rice, losses are the highest during harvest (11.1 percent) and cleaning (6.0 percent), and the lowest for storage (1.4 percent), although for the latter, the breadth of the range of plausible values (0.1 percent to 4.1 percent) makes it impossible to draw any clear conclusions. Loss patterns for millet and sorghum are similar, with harvest losses being the greatest (6.4 percent and 4.3 percent, respectively) and much lower losses for the other post-production stages.

The calculated confidence intervals indicate that the gap between harvest losses for rice and maize is not statistically significant. The difference between harvest losses and losses during other operations, with the former being higher than the latter, is statistically significant for all crops, reaching the level of 10 percent. The confidence bands for rice are wider than for the other crops, as shown earlier in this report for yields (declared and measured) and declared losses (cf. previous section).
Table 12. Crop losses by operation (objective measurements).

<table>
<thead>
<tr>
<th>Averages</th>
<th></th>
</tr>
</thead>
</table>
| Harvest  | Threshing | Cleaning | Storage
| Millet   | 6.4        | 1.7      | 2.1     | 1.0     | [5.8 – 7.9] | [1.6 – 2.1] | [1.8 – 2.6] | [0.7 – 1.6] |
| Maize    | 9.0        | 3.4      | 4.5     | 4.1     | [8.6 – 10] | [3.3 – 3.8] | [4.3 – 4.8] | [3.3 – 5.8] |
| Rice     | 11.1       | 2.7      | 6.0     | 1.4     | [8.7 – 16.9] | [2.4 – 4.0] | [4.2 – 8.6] | [0.1 – 4.1] |
| Sorghum  | 4.3        | 1.2      | 1.2     | 2.5     | [3.9 - 5.3] | [1.0 - 2.1] | [1.0 – 1.4] | [1.7 – 4.2] |

The profile of storage losses occurring over time indicate an increase in the percentage losses noted during the second visit (approximately one month after the first visit) for the four crops. The measurements made on grain samples taken at the third and last visit indicate slight decreases in the losses registered for maize, rice and millet. However, given the standard deviations and confidence bands for storage losses, these differences are not statistically significant. Conversely, storage losses increase during the last visit; however, in this case too, the variations in the estimates do not allow unequivocal conclusions to be drawn. Grain losses during storage are normally expected to increase with storage duration. However, as indicated by Stathers, Lambell and Mvumi (2013), this effect generally kicks in after the third month, when insect infestation starts to reach seriously damaging intensities. This may explain why the changes in storage losses over time estimated here are not significant.

Figure 9. Storage losses over time (% all districts).

---

14 At the third and last visit.
5.4.3. Farmers’ declarations versus objective measurements

In this study, field measurements have led to higher loss estimates than farmer reporting (see figure 10). For maize, losses at harvest are estimated at 9.0 percent and 4.1 percent, respectively. Similar observations can be made for the other crops and other post-production stages (threshing/shelling, cleaning, and storage). These differences are statistically significant, as evidenced by the non-overlapping confidence bands.

The literature comparing these two loss estimation techniques is scarce. Among the few studies available, the *Post-Harvest Loss Assessment Survey in Malawi* conducted in 2011\(^\text{16}\) has confirmed this finding: maize losses are estimated at 9.7 percent when using measurement techniques, well above the estimate of 1.5–2.2 percent obtained through farmers’ declarations.

*Figure 10. Maize losses: farmer-reported versus objective measurements, by operation (in %, all districts).*

---

5.4.4. General discussion and further analysis

The estimates are in line with evidence from the recent literature

Ratinger (2013)\textsuperscript{17} compiled information from different studies for different countries and regions. For Ghana, he provides HPHL estimates for maize in the range of 7 percent to 14 percent. A survey conducted in 2013/2014 in Malawi\textsuperscript{18} estimated HPHL for maize (as reported by farmers) at 10.7 percent. Our estimate of 9.4 percent for declared losses is consistent with these figures.

For rice, Ratinger (2013) indicates total losses of 11 percent in Uganda, close to the estimate provided in this report on the basis of objective measurements. Appiah \textit{et al.} (2011) estimate total PHL (at harvest, threshing and drying) for rice to be between 4.6 percent and 17.9 percent, depending on the rice variety and production practices. The same authors estimate harvest losses to fall within the range of 3 percent to 12 percent. Our estimate based on field measurements (11.1 percent) is in the high end of this range, while our farmer-based estimate (2.9 percent) is on the lower end.

It is also interesting to compare the results of the present study with those obtained through approaches that are not fully based on sample surveys, farmer interviews or field measurements. APHLIS, for example, is a model producing calculated PHL estimates for food crops across sub-Saharan Africa, currently covering major cereal crops in 38 countries of the region. These estimates tend to be higher than the declaration-based estimates presented in this study. For example, in the case of maize, APHLIS estimates PHL at 18 percent in Ghana, well above the 9.4 percent declaration-based estimate of the present study. When comparing the APHLIS results for maize with our measurement-based estimates, the differences are less marked.

Given the nature of the methods used in the APHLIS model, the context-specificity of some of its parameters, and the lack of transparency in the calculation process, it is difficult to assess the quality of these estimates and understand the causes of any difference. The stability over time of some of the estimates (maize losses for Ghana are estimated at 18 percent for 2003 up to 2013, for example) may seem unrealistic given the dependence of losses on

\textsuperscript{17} Ratinger, T. 2013. \textit{Food Losses in the Selected Food Supply Chains}. Selected Paper prepared for presentation at the 140\textsuperscript{th} Seminar of the European Association of Agricultural Economists, 13–15 December. Perugia, Italy.

volatile climatic conditions (humidity, rainfall, etc.) and changing farming practices.

**Weak apparent correlation between observed and declared losses**

The scatter plots and regression lines between measured and declared losses for maize (figure 11) point to a weak apparent correlation between these two assessment methods. It is only in the case of threshing losses that a significant and positive correlation can be observed. This analysis would require further refinement to reach conclusions on the nature and extent of the relationship between observed and declared losses, for example by controlling in a multiple regression for factors that may be correlated with losses, such as farmers’ education, threshing method, length of experience in agriculture, farm size, type of storage facility and seed variety. Figure 11 also illustrates the systematic differences between farmer-reported and measured losses, the latter being in most cases higher than the former.

Finally, the scatter plots indicate a greater variability between declared losses and measured losses when the declared losses are low. This could be explained by the fact that: farmers may have little knowledge about the actual losses they experience (assuming that the measured losses are closer to the real losses experienced by the farmer); and/or that they consider there is a loss only when losses are higher than usual – that is, when there is an exceptional event such as a major pest infestation, an intense rainfall or large crop damages due to a mechanical failure.
Figure 11. Crop losses: farmers’ declarations versus objective measurements (%, maize, all districts).
Conclusion

The objective of this report was to present the field test carried out in Ghana on the measurement of HPHL, comparing estimates based on farmers’ declarations with those based on field measurements. This report presented the methodological design of the study, the organization of the data collection activities, the calculation procedures and the main results.

The estimates presented and discussed in this report are broadly in line with those available in the recent literature for similar grain crops, regions and countries. They provide additional support to the fact that estimates based on objective measurements tend to be consistently higher than farmer-based estimates, a finding encountered in several studies. This is true for production and yields but also for crop losses, a topic that has benefited from far less attention in the debate on objective measurements versus farmer declarations.

The present study provides new and useful evidence on crop losses for harvest and post-harvest operations for four of Ghana’s major crops in two districts of the country. This evidence points to a weak apparent correlation between measured and declared losses; however, this relationship should be further analysed by including other explanatory variables.

Beyond the data and indicators provided by this pilot survey, useful insights can be learned from the data collection and estimation processes: measuring crop losses on the farm is a complex undertaking, for respondents and enumerators alike. It requires the use of skilled and experienced data collection teams and well-defined questionnaires that customized to the local context and reflect actual farming practices. A thorough training and pre-testing of data collection tools is also necessary, especially when objective measurements are envisaged. Indeed, as their complexity is greater than that of standard crop-cutting exercises, they require adequate training and the provision of appropriate measurement tools, such as weighting scales with the sufficient level of precision or spears to select grain samples. For example, during this pilot survey, the inadequacy of some weighing scales used caused delays in the field activities and affected some of the measurements. This pilot survey also highlighted the need to adapt the data collection and measurement approach to better account for mechanized practices for harvest and post-harvest operations.
Given the technical challenges related to the measurement of crop losses and the differences in the results between objective and declarative procedures, it seems appropriate to combine the two approaches. Further work is necessary to understand how measured and declared losses can be combined into a sound modelling framework, making use of variables that may have a bearing on losses, such as harvesting practices, type of storage facilities, crop variety and exogenous factors such as climatic conditions. A larger sample of farms and measurements is needed to conceive and test such econometric models.
Annexes

Annex 1: Summary tables

Table 1. Average crop yields and losses by operation (farmers’ declarations).

<table>
<thead>
<tr>
<th>District</th>
<th>Crop</th>
<th>kg/ha</th>
<th>In %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Y</td>
<td>HL</td>
</tr>
<tr>
<td>Kintampo</td>
<td>Maize</td>
<td>984</td>
<td>5.3</td>
</tr>
<tr>
<td></td>
<td>Millet</td>
<td>1 416</td>
<td>4.3</td>
</tr>
<tr>
<td></td>
<td>Rice</td>
<td>1 513</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>Sorghum</td>
<td>1 584</td>
<td>2.8</td>
</tr>
<tr>
<td>Sawla</td>
<td>Maize</td>
<td>936</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Millet</td>
<td>537</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>Rice</td>
<td>1 116</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>Sorghum</td>
<td>790</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Table 2. Average crop yields and losses by operation (objective measurements).

<table>
<thead>
<tr>
<th>District</th>
<th>Crop</th>
<th>kg/ha</th>
<th>In %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Y</td>
<td>HL</td>
</tr>
<tr>
<td>Kintampo</td>
<td>Maize</td>
<td>2 270</td>
<td>7.3</td>
</tr>
<tr>
<td></td>
<td>Millet</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td></td>
<td>Rice</td>
<td>2 182</td>
<td>11.4</td>
</tr>
<tr>
<td></td>
<td>Sorghum</td>
<td>1 065</td>
<td>6.5</td>
</tr>
<tr>
<td>Sawla</td>
<td>Maize</td>
<td>2 282</td>
<td>33.2</td>
</tr>
<tr>
<td></td>
<td>Millet</td>
<td>1 068</td>
<td>6.4</td>
</tr>
<tr>
<td></td>
<td>Rice</td>
<td>1 176</td>
<td>10.0</td>
</tr>
<tr>
<td></td>
<td>Sorghum</td>
<td>1 486</td>
<td>3.8</td>
</tr>
</tbody>
</table>
Annex 2: Example of questionnaires

FH3 – Production and losses: estimation by enquiry

FH6 – Production and losses: estimation by objective measurements
Annex 3: Estimation of variances, standard deviations and confidence intervals

Variance and standard deviation

For a given district and crop, the average parameters are estimated as ratios of district-level estimates:

\[ \hat{l} = \frac{\hat{L}}{\hat{Q}} = \sum_{j \in EA} \theta_j(q) . \hat{t}_j \]

The estimated variance and standard deviation of this ratio can be approximated respectively by:

\[ \hat{V}(l) = \left( \frac{\hat{Q}}{\hat{L}} \right)^2 \left[ \hat{V}(\hat{L}) + \hat{V}(\hat{Q}) - 2 \frac{\text{COV}(\hat{L}, \hat{Q})}{\hat{L} \cdot \hat{Q}} \right] \text{ and } \hat{SD}(l) = \sqrt{\hat{V}(l)} \]

where:

\[ \hat{V}(\hat{L}) = \frac{1}{n(1 - n)} \sum_j \left( \frac{\hat{L}_j - \hat{L}}{n} \right)^2 \]

\[ \hat{V}(\hat{Q}) = \frac{1}{n(1 - n)} \sum_j \left( \frac{\hat{Q}_j - \hat{Q}}{n} \right)^2 \]

\[ \text{COV}(\hat{L}, \hat{Q}) = \frac{1}{n(1 - n)} \sum_j \left( \frac{\hat{L}_j - \hat{L}}{n} \right) \left( \frac{\hat{Q}_j - \hat{Q}}{n} \right) \]

with \( n \) being the number of EAs with the selected crop in each district.
The average estimate over the two districts (referred to as $\hat{l}$ here for convenience) is: $\hat{l} = \theta_1 \hat{l}_1 + \theta_2 \hat{l}_2$, where $\theta_1 = \frac{\hat{q}_1}{\hat{q}_1 + \hat{q}_2}$ and $\theta_2 = \frac{\hat{q}_2}{\hat{q}_1 + \hat{q}_2}$. Its estimated variance is therefore given by (excluding the covariance term): $\hat{\nu}(\hat{l}) = \theta_1^2 \hat{\nu}(\hat{l}_1) + \theta_2^2 \hat{\nu}(\hat{l}_2)$.

**Confidence intervals**

Confidence intervals for a given risk level $\alpha$ have been constructed using the following formula:

$$IC_\alpha(\hat{l}) = \left[\hat{l} - F_{\alpha/2}^L \cdot \text{SD}(\hat{l}) ; \hat{l} + F_{1-\alpha/2}^L \cdot \text{SD}(\hat{l})\right]$$

where $F_p^L$ is the $p$-quantile corresponding to the probability distribution $L$ of the standardized transformation of $\hat{l}$. The probability distribution $L$ was estimated using a Gaussian smoothing window (or kernel). The probability distributions were estimated for each parameter (yields, harvest losses, threshing losses, etc.), pooling data for all EAs and crops within a district to obtain sufficiently robust estimates.