50x2030
DATA-SMART AGRICULTURE

RESEARCH ON THE MEASUREMENT OF HARVEST AND POST-HARVEST LOSSES:

On-Farm Storage Loss Model to Improve National Estimates and Reduce Data Collection Costs

June 2025

50x2030 WORKING PAPER SERIES

Summary

As much as awareness increased on the importance of post-harvest losses and the need for sound data for decision-making, the costs associated with data collection becomes an inhibiting factor for regular national surveys, especially for countries with limited resources. Resource-optimizing strategies for generating reliable data become pivotal to reduce data collection costs. In this regard, the 50x2030 Initiative to Close the Agricultural Data Gap promotes methodological innovations that improve data collection and reduce survey costs in the agriculture sector. As part of the 50x2030 research, the food loss research component explores approaches that allow to combine farm or rural household surveys with statistical modelling methods to improve estimates and reduce data collection needs. This can be especially relevant for measuring on-farm losses, given that its measurement can be complex and often requires resource-intensive methods. One of the most relevant on-farm loss points, but at the same time most demanding in terms of measurement, are grain storage losses.

This research aims to develop and test a modelling approach to estimate grain storage losses on the farm, building on regularly conducted national agriculture surveys. The modelling approach is meant to allow countries to apply occasional in-depth survey modules based on physical measurement methods of storage losses to achieve data quality, while using a modelling approach in the subsequent survey rounds of the agriculture survey for indirect estimations. This approach can be relevant for those countries with a specific interest and policy priority on on-farm storage losses, for which the quality and accuracy of the storage loss estimates is of importance, as well as insights on the main drivers and characteristics of the storage activities collected with the in-depth survey module.

For this purpose, the research builds on several methodological developments. In a first step, a detailed conceptual framework of grain storage losses is drafted from relevant literature about storage characteristics, losses and their causes. Based on the conceptual framework, the data needs are derived, and the in-depth storage loss survey is designed. The approach is then pilot tested in Mali, collecting the data on storage losses and on a wide set of explanatory variables in the selected pilot region, focusing here on maize and millet Based on the data collected, modelling approaches to estimate storage losses and to identify explanatory variables are tested. As an important component, the survey data is complemented by data from the national farm survey, here EAC survey 2022/2023, as well as other data sources providing weather information. The research concluded with the baseline model to estimate grain storage losses for the first survey round, specifying the main explanatory variables and establishing the model function.

As a result, it can be highlighted that the in-depth storage loss module, together with indicators on the farm characteristics and weather variables, produces a wide range of significant explanatory variables for the grain storage losses estimated in Mali on maize and millet Grain storage losses, obtained from laboratory analysis of grain samples from three visits, were estimated at 2.5% for maize and 0.6% for millet, which is in line with the general literature. Nevertheless, these values might also be underestimated, given that they measure mainly losses due to pest infestations. Though these are considered the mayor cause of loss during storage, the used methods exclude to some extend rodent

attacks and removed and discarded grains (i.e. infested, damaged packaging, etc). A major advantage of the methods is reflected in the quality of the obtained estimates, with relatively low standard deviations compared to declaration-based storage losses.

Based on the obtained estimates and explanatory variables, the exercise generated baseline models with sufficiently good model specifications. The models are built on Generalized Estimating Equations (GEE) using the natural logarithm as the link function and the Poisson distribution family. Both baseline models passed the model specification test, with stronger results for maize compared to millet A major limitation has been the sample size, whereby the number of explanatory variables in the models was successively reduced. Being a specific objective, it was possible to build the models using not only variables collected in the in-depth storage loss module but making use of indicators that are already collected in the agriculture farm survey in Mali, EAC survey 2022/2023, and from auxiliary data sources. This helps to reduce the number of variables countries would need to collect additionally for estimating storage grain losses using their national farm survey. Some of the selected variables relate to the characteristics of the storage facility (i.e. type of walls, raised rackets, packaging material), or to drying methods (i.e. floor drying).

Some challenges were faced in terms of the complex field operation of applying physical measurements to produce better quality estimates of storage grain losses, requiring laboratory analysis and several visits to the farm. Due to the large number of 1290 grain samples taken, operational challenges occurred to deliver and analyse these in a relatively short time. In some cases, farmers refused to provide access to the storage facility or grains stored in hermetic bags to avoid contamination. These challenges can be mitigated through institutional capacity building to carry out the analysis of grain samples in the field and the generation of visual scales that could replace the laboratory analysis. The provision of compensations for the farmers to participate in the survey might be needed, considering the potential losses incurred due to the grain sample extraction. Another challenge has been the timing of the survey, with the first visit starting few months after harvesting and the last visit being three months away from the reported end of the storage period. A proper planning exercise should be done prior to the survey implementation to obtain information on the storage period, as well as storage characteristics.

This research provides the first elements, but further research is required to be able to estimate grain storage losses indirectly in years when the national farm survey does not include the storage loss module. To complete the modelling approach, a second in-depth storage loss survey is required, thereby allowing to incorporate interannual variations in the modelling function, making it possible to estimate storage losses for intermediate years. The model can then be further calibrated with each subsequent application of the in-depth storage loss module.

Once established, the periodicity of storage loss measurements could be considerably reduced, eventually to every 3rd to 4th survey round or every 6-8 years in national farm surveys. In turn, countries would benefit from improved data quality of storage losses and of a wide set of indicators on the storage characteristics, activities and trends, that are relevant for decision-making.

Contents

List of Figures

List of Tables

Abbreviations

50X2030	The 50x2030 Initiative for Data smart Agriculture
AFRISTAT	Planning and Statistics Unit of the Rural Development Sector
APHLIS	African Post-Harvest Losses Information System
C3S	Copernicus Climate Change Service
CAPI	Computer Assisted Personal Interviews
CART	Correlation Analysis, and Regression Trees
CIMAT	Centro de Investigación en Matemáticas (Center for Research in Mathematics)
CPS/SDR	Cellule de Planification et de Statistique (CPS) du Ministère de l'Agriculture du Mali (Planning and Statistics Unit of the Ministry for Agriculture in Mali)
EAC	Enquête Agricole de Conjoncture (Agricultural Economic Survey).
ERA5	The fifth generation European Centre for Medium-Range Weather Forecasts atmospheric reanalysis of the global climate
INSTAT	Institut National De La Statistique du Mali (Mali National Institute of Statistics)
FAO	Food and Agriculture Organization of the United Nations
FLW	Food Loss and Waste
GEE	Generalized Estimation Equation
GSARS	Global Strategy for Agriculture and Rural Statistics
LABOSEM	National Seed Laboratory
SDG	Sustainable Development Goals
SSA	Sub-Saharan Africa

Acknowledgement

This paper was produced with financial support from the 50x2030 Initiative to Close the Agricultural Data Gap, a multi-partner program that seeks to bridge the global agricultural data gap by transforming data systems in 50 countries in Africa, Asia, the Middle East and Latin America by 2030.

With the technical support from Food and Agriculture Organization of the United Nations and the World Bank. It was drafted by Daniela Ruehl (FAO), Ignacio Mendez-Gomez-Humaran (Centro de Investigación en Matemáticas, CIMAT), Sharon Masakhwe Mayienga (FAO), and Marco Constantini (World Bank), with guidance from Marco Tiberti (World Bank) and Carola Fabi (FAO).

The authors want to thank Franck Cachia (FAO) for formulating the research component and the research idea that led into this research and Mbaye Kebe for his technical guidance throughout the research. They also acknowledge the support of FAO and the World Bank in providing resources that enabled the process to be a success.

A special thanks go to AFRISTAT, in especially Dine Djabar Adechian and Assitan Traoré, as well as to the Planning and Statistics Unit of the Ministry of Agriculture of Mali (Cellule de Planification et de Statistique (CPS) du Ministère de l'Agriculture du Mali), being our key partners in Mali and without whom the data collection could not have been successfully achieved.

1 Introduction

1.1 Background and rationale

Reduction of food loss and waste, being a challenge of global concern, has been established as one of the Sustainable Development Goals (SDG) under target 12.3. While there is growing awareness of the importance of food loss reduction at the political level, official post-harvest loss data for informing policymaking and reporting on the SDG Indicator is scarce. Notably, food loss measurement is complex, given it happens in several stages along the food supply chains, resulting in a burdensome challenge for countries. Therefore, the first starting points for countries is to measure losses on the farm, where losses can be critical in terms of quantities getting lost and its impact on food security, while data collection can be leveraged on existing on-farm data collection systems.

The 50x2030 Initiative to Close the Agricultural Data Gap[1], in its effort to promote and implement national agriculture surveys, recommends and supports the integration of food loss modules. A specific on-farm food loss research component has been carried out with the Food and Agriculture Organization of the United Nations (FAO), and in collaboration with the World Bank to identify methodological innovations that can help improving data quality and reducing data collection costs of food loss measurement in national agriculture, farm or household surveys. This paper presents the results of the third and last research component[2], focusing on on-farm storage losses and the possibility to use modelling approaches to reduce the cost and frequency of data collection, while improving the quality of estimates when storage loss data are collected through farm surveys.

On-farm storage losses have been subject to public discussion for several decades, especially for those products stored for longer periods on the farm as cereals, pulses, roots and tubers. These are relevant staple crops in most countries in Sub-Saharan Africa, making their losses critical for food security and subject to national policy priorities. The literature indicates that on-farm storage losses of grains can be as much as 50%–60% if the grain is not well protected, while the use of improved storage methods can reduce these losses to as low as 1%–2% (Kumar et al., 2017). Survey-based and model-based grain loss estimates for on-farm storage seem to range in general between 1-7% (Abdoulaye et al., 2016; Hodges et

---

[1] The 50x2030 Initiative to Close the Agricultural Data Gao is a 10-year multi-agency partnership that seeks to bridge the agricultural data gap by transforming data systems in 50 countries in Africa, Asia, the Middle East and Latin America by 2030. It is implemented through a unique partnership between the World Bank, Food and Agriculture Organization of the United Nations (FAO) and International Fund for Agricultural Development (IFAD). More information can be found here https://www.50x2030.org/.

[2] The following outputs have been produced for this research project on post-harvest losses funded by 50x2030 Initiative to Close the Agricultural Data Gap:

Ruehl, Mendez-Gomez-Humaran, Tiberti. 2024. *Combining food loss modelling approaches with farm surveys to improve on-farm loss estimates and reduce data collection costs*. 50x20230 Working Paper Series.

Mayienga, Cachia. 2021. *Minimum Losses by Commodity and Region: Insights from the Literature.* 20x2030 Working Paper Series.

al. 2011; FAO/ESS, 2023; FAO, 2017). Considering total grain losses, storage accounts for approximately 20 to 30 percent of all the losses in the value chains (Manandhar et al., 2018).

But measurement of losses during storage is complex. Often, storage losses are due to insect pests and rodent attacks, but also to deterioration driven by microorganisms, and to some extent spillage. The quantities of grains are thereby reduced, though it is not always visible or observable. Adding on this, losses occur throughout a prolonged period, with its cumulated quantities being difficult to sum up to a total. In some cases, the crops are stored until the next harvesting period, at the same time, grain is often removed or added to the storage, or it is moved between different storage facilities on the farm. Farmers might occasionally remove and discard damaged and infested grains from the storage. This can considerably limit the recall capacity of the farmers of their total storage losses. The timing of when to declare storage losses is also a challenge in national agriculture surveys. Most agriculture surveys are implemented close to the harvesting period of the main crops, when the crop is recently stored and storage losses low, thus estimates would underestimate these. Additional visits might be needed, or farmers would need to declare storage losses of their former harvesting period, which can further affect their recall capacity.

Given these challenges, there are limitations to use declaration-based methods to measure storage losses. Therefore, research efforts were made on the design and use of physical measurement methods to generate better quality data, which date back to the 1970s and are summarized by Boxall (1986). These are based on analysing grain samples from the storage, with various assessment methods being developed and tested to measure to what extent the grain is damaged and how much, in consequence, has been lost in terms of weight during the storage period. While these methods are likely to generate better data quality when it comes to grain storage losses, it can considerably increase the cost of data collection in a national survey. Therefore, their application in national farm surveys is recommended only in combination with strategies to reduce data collection costs. The Guidelines to measure post-harvest losses of grains by the Global Strategy to Improve Agricultural and Rural Statics outline how to use sub-sampling strategies of farms where to apply the physical loss measurement (GSARS, 2018). The 50x2030 Initiative provides orientation on how to design modular agricultural surveys that allow to rotate modules so that some modules, such as the food loss module, are collected with a reduced frequency than the survey (50x2030 Initiative, 2021).

In this context, the overall objective of this research activity is to explore if these cost reduction strategies can be further strengthened by using statistical models and estimate grain storage losses indirectly when the storage loss module is not applied, allowing to further reduce the frequency of collecting storage loss data.

For doing so, a detailed in-depth storage loss survey module, collected as part of the national farm survey, aims at introducing physical measurement approaches for storage grain losses, while providing the required set of explanatory variables and drivers, such as the type of storage facility and packaging, the application of pest protection products, the duration of storage, and the general agro-climatic conditions. The modelling approach is then used to choose a relevant set of explanatory variables collected in the in-depth survey and to establish a baseline model for the estimation of storage grain losses. With a second in-depth storage loss survey, interannual variations of storage grain losses could be factored in, allowing for intermediate indirect storage loss estimates based on explanatory variables collected in the national farm survey for those years the storage loss module is not applied. The storage loss module might then be applied every 3rd or 4th survey round only, with indirect estimations being produced instead.

Therefore, it represents a strategy for countries to choose to invest into better quality of storage loss estimates but on a considerable reduced frequency, leveraging on the national agriculture survey to produce indirect storage loss estimates. It brings the advantage of providing a whole set of relevant explanatory variables and characteristics of the storage activity that can be relevant for decision-making and policy design. Therefore, this approach can be of specific relevance and interest for those countries where on-farm grain storage and corresponding losses have been identified as a key policy priority and for which improved evidence-based is needed.

1.2 Objectives and scope of the research

General objective

The main long-term objective of this research is to develop and test a statistical modelling approach that allows to estimate grain storage losses on the farm indirectly, based on a set of explanatory variables collected in the farm survey. While this is the ultimate objective, this research focuses on designing the in-depth storage loss survey module and implement it in one country, in order to establish from there the baseline model to estimate storage grain losses. This baseline model would cover the first survey round, putting an emphasis on screening and identifying the main explanatory variables and the model structure.

Specific objectives

The specific objectives consist of developing and field testing the methodological elements required for an in-depth storage loss module in national farm surveys and modelling of on-farm storage losses. More specifically, these consist of:

1. To summarize, based on existing literature, the conceptual framework of on-farm grain storage losses, its measurement and the main causing factors.
2. To design the corresponding in-depth storage losses module that uses physical measurement methods and collects a diverse set of causal factors.
3. To field-test the in-depth storage loss survey module in one pilot country as a sub-sample of the national farm survey.
4. To design and test modelling approaches to identify key explanatory variables and establish the baseline model to estimate storage grain losses, using the available indicators from the national farm survey, the in-depth storage loss module and other data sources on its driving factors.

Area of application

The strategy for using modelling approaches as part of national farm surveys is visualized in Figure 1. As mentioned, the starting point is the integration of the in-depth storage loss survey module in the national farm survey, using physical measurement methods that are meant to generate higher-quality data compared to farmer-reported, declaration-based estimates. The storage loss module, including a detailed set of information on the storage activity and characteristics, is applied to a sub-sample of the national farm survey. The integration into the national farm survey then allows to make use of the general set of farm indicators usually collected in the storage loss modelling, such as general farm characteristics.

Based on the comprehensive set of potential explanatory variables obtained from the in-depth storage loss module, the modeling approach will identify which variables are most relevant for explaining and estimating grain storage losses. These variables can then be integrated into the core modules of the national farm survey. Based on a second survey round, this would allow for indirect estimation of storage losses of intermediate as well as future survey rounds when the full storage loss module is not applied.

The in-depth module, implemented in year 1 (see Figure 1) also helps distinguish between *structural variables*—which change only in the medium to long term (e.g., storage infrastructure, drying equipment or methods, agro-ecological zone)—and *short-term drivers* that influence year-to-year fluctuations in storage losses. The latter include factors like weather conditions, pest incidence, the quantity of grain stored, and storage duration. These short-term variables are key inputs into the modeling approach, especially when having the second survey round to account for interannual variations.

In upcoming years, the national farm survey will continue without the dedicated storage loss module but will collect data on the key explanatory variables. Structural variables—such as storage infrastructure—can be carried forward from the most recent in-depth module, as they are assumed to remain stable over several years. Short-term variables and key explanatory variables should be included in the annual survey to allow for model-based estimation of storage losses. Some of these variables (e.g. pest incidence, drying activity, quantity stored, and storage period) can be collected through the survey itself, while others (e.g. weather data) may be sourced from auxiliary data systems. It is likely that not all explanatory variables will be consistently available. Therefore, a *minimum required dataset* must be defined, focusing on the most critical variables. This minimum set will be cross validated with data availability in each country. Certain information—such as weather data, pest outbreaks, and price trends—may already be regularly produced by public institutions and can be integrated accordingly.

Figure 1 illustrates this strategy using the example of Mali's national farm survey, the *Enquête Agricole de Conjoncture* (EAC), implemented annually by the Planning and Statistics Unit of the Rural Development Sector (CPS) in the Ministry of Agriculture. In years 1 and 4, the full in-depth storage loss module is applied using physical measurement methods. In years 2 and 3, storage losses are estimated indirectly using a model based on explanatory variables collected in the regular farm survey and auxiliary data sources such as weather data. In consecutive years 5 and 6, storage losses will be estimated indirectly based on two applications of the full in-depth storage loss module producing more efficient and reliable loss estimates.

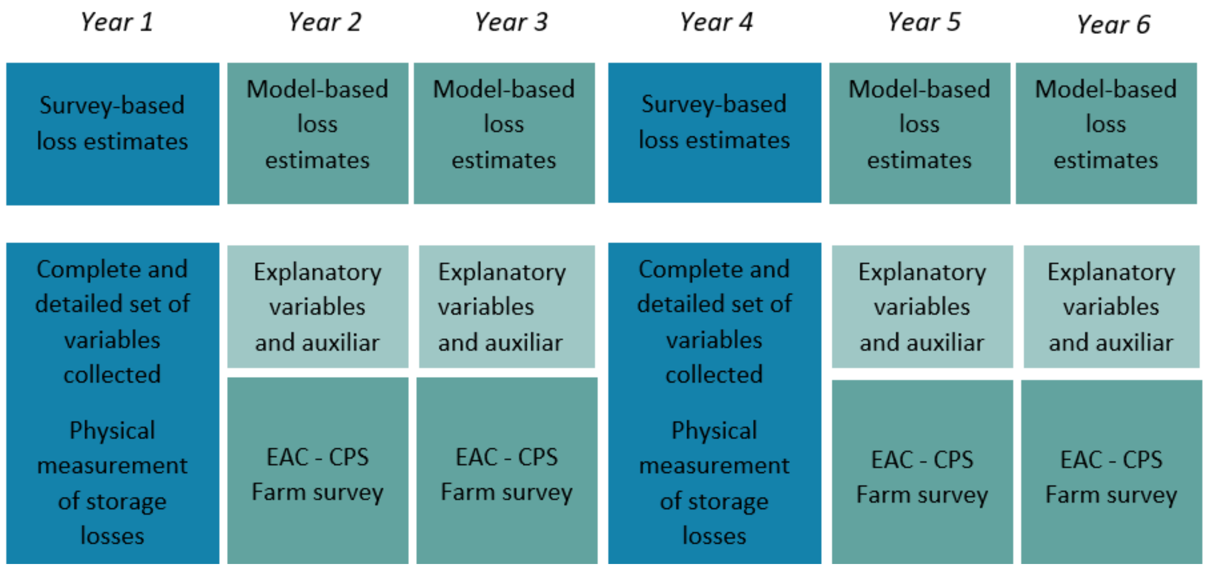

Figure 1: Storage loss models to indirectly estimate loss based on farm surveys

Scope of this research

While Figure 1 illustrates the potential for improving the quality of grain storage loss estimates and reducing data collection costs through model-based approaches in future survey rounds, this research represents only the first step towards that application, focusing on year 1 of Figure 1. The in-depth storage loss survey will be designed and collected from a sub-sample of the EAC survey 2022/2023. Based on the obtained storage loss estimates and possible explanatory variables, a procedure will be established for measuring and modelling storage losses during one survey implementation (one harvest period), which represents the baseline (cross-sectional study). This baseline model defines the explanatory variables that serve as predictors, as well as the model function that accounts for storage losses occurring throughout the storage period and covered by several visits and storage loss observations.

To establish intermediate estimates of grain storage losses based on the relevant predictors collected in the forthcoming survey rounds (Year 2 and 3), this research needs to be further developed to consider temporal variability in the modelling approach by accounting for interannual variations in losses. For doing so, a second implementation of the in-depth storage loss survey in year 4 is needed. Once two survey rounds are completed, indirect storage loss estimates can be produced from year 5 and 6 using two consecutive applications of the in-depth storage loss module.

1.3 Pilot country Mali and institutional arrangement

Mali was chosen as the pilot country to test the approach. It is one of the countries that has received technical assistance from under the Global Strategy for Agricultural and Rural Statistics (GSARS) on how to measure harvest and post-harvest losses on-farm. A module on measurement of food losses on-farm was therefore included in their annual agricultural questionnaires, the Enquête Agricole de Conjoncture (EAC) and implemented, while storage losses are not included in the measurement. This provides the opportunity to further explore how to integrate storage loss measurement in the national agriculture survey, in close collaboration with the Planning and Statistics Unit of the Rural Development Sector (CPS/SDR) of the Ministry of Agriculture.

Additionally, Mali is the headquarter of the Economic and Statistical Observatory of Sub-Saharan Africa (AFRISTAT) considered as a strategic partner for implementing methodological development activities and strengthening the statistical capacities of African countries. AFRISTAT collaborated in the research component as local implementation partner for the field testing, while they signed a memorandum of understanding with CPS/SDR for a joined implementation.

To ensure the involvement of all relevant national institutions, a technical committee was established comprising the Economic and Statistical Observatory of Sub-Saharan Africa (AFRISTAT), the Planning and Statistics Unit of the Rural Development Sector (CPS/SDR), the National Institute of Statistics (INSTAT), and the National Seed Laboratory (LABOSEM), to support the implementation of various activities related to all aspects of the investigation.

2. Conceptual framework of grain storage losses and its modelling

2.1 Overview of grain storage losses

As a starting point, storage losses are briefly conceptualized with regards to its measurement and the main causing factors. This allows to better identify the data needs and the measurement methods. Focusing on quantitative food losses, as defined in the SDG 12.3.1 target, the grain gets physically lost, not being available for consumption anymore, or it is damaged to a point where it is not fit for human consumption and discarded. The Guidelines on the measurement of harvest and post-harvest losses of cereals by the Global Strategy to Improve Agriculture and Rural Statistics (GSARS, 2018) distinguish several causes for grain losses, biological and microbiological, biochemical and chemical, environmental and climatic, and mechanical or technical. These are also considered the primary causes of storage losses. For its measurement, this research briefly describes those resulting in a quantitative loss during storage, differentiating them with a view on how it is measured, as summarized by Boxall (1986). These are:

1. **Grain losses due to insect pests during storage:** One of the major challenges for storing grains are insect pests attacking the grains while they are stored. When being attacked, the grains lose weight and remain often affected by presenting holes, breaking, and losing their natural protection. The whole grain is rarely eaten, leaving some remaining parts back in storage. These are observable, but difficult to quantify. Various methods to measure these based on physical measurement have been developed, which are further described in the next section. The quantity of grain affected by insect attacks usually increases with the time the grain is in the storage, especially when no measures are taken during storage to combat them. In the context of developing countries, these are often the most relevant cause of grain losses during storage (Abdoulaye et al., 2016).

2. **Grain is affected by microorganisms, particularly fungi/moulds, during storage (deterioration):** Another reason why grain is lost during storage is due to growth of micro-organisms, particularly moulds. This can be caused by high level of humidity due to poor drying practices, damages of the grain from harvest and post-harvest handling or level of contamination and foreign materials in the grain. The overall climatic conditions in the area, during harvest and post-harvest handling and during the storage period can therefore be key if the grain is not properly protected. The deterioration is only to some extent observable by the farmer. Similarly to insect pests, the infested grains lose weight, which is difficult to observe and estimate. These can be similarly measured as losses by insect pests. Nevertheless, in some cases, the farmers might spot the infestation and remove the infested area from storage to avoid further spreading, especially when stored without being packed. These interventions can be very occasional and depend on the circumstances and local practices. If no action is taken, the quantities affected by the infestation will increase with the time the grain is stored. Infested grains might also be discarded after storage, when it is rejected by buyers.

3. **Grain is lost to rodents and birds during storage (vertebrate pests):** Stored grain can also diminish and get lost because of rodents and birds consuming the grains. These are often less

observable than insect and microorganism infestation and therefore difficult for the farmer to quantify. For some grains, rodents might only eat the centre part of the grains, in other cases, remove the whole grain. Often, they open bags and thereby indirectly affect all grains in the bag. Some physical measurement methods and approaches were developed and tested, depending if the grain is stored on the cob or head, or threshed. Given that storage conditions and protection of the grain are often unchanged throughout the storage period, the quantity of grain lost to rodents and birds will increase with the time the grain is in the storage.

4. **Grains are spilled during storage handling processes (loading, removing, etc):** Grains are often spilled during handling operations such as loading into storage, removing from storage, or transferring between containers. Spillage can occur due to damaged packaging materials or improper handling — whether manual or mechanical. These losses are visible, as the spilled grains typically remain on the ground or inside storage units. Such losses are often linked to inadequate packaging and poor handling practices. They may be more common when large quantities are handled, particularly when machinery is used, as this can lead to less careful handling.

5. **Humidity weight loss of stored grains not considered quantitative food loss**: The presented loss points lead to losses of the total weight of the stored grains and are all considered quantitative grain losses. Nevertheless, not all weight losses are related to a physical loss of the grain. During storage, the grains´ moisture content usually decreases, whereby the same amount of grains weigh less than before. Although the weight is lower, the reduction of moisture content is not considered a food loss and must be excluded from the assessment. For farmers, though, they are still relevant given that the price of grain depends on the total weight.

6. **Grain storage losses and the time factor:** An important characteristic of storage losses is the duration of storage and its impact on grain losses. This relationship is subject to several research studies, whereby it can be seen that percentage losses tend to increase with storage duration if storage conditions are inappropriate to protect the grain (Mendesil et al., 2022; Mlambo et al., 2017, Freitas et al., 2016). Total storage losses occur during a prolonged period while the produce is stored, and how much was lost during storage can only be known once the produce is not in the storage anymore.

2.2 Causes and possible driving factor of grain storage losses

Based on the description of the primary causes of grain losses in storage, a literature review has been conducted to identify articles that outline the underlying drivers of these types of storage losses, also referred to as secondary causes of storage losses. Drivers are those factors that enable and influence the extent to which the primary causes can expand and cause the loss of grains. These are key as explanatory factors and will define the data to be collected for the storage loss modelling work. The search for articles was carried out on the FAO Food Loss and Waste Database[3] and complemented by articles from google

[3] FAO Food Loss and Waste Database: https://www.fao.org/platform-food-loss-waste/flw-data/en/

scholar and the International System for Agricultural Science and Technology (AGRIS) repository[4]. The African Post-harvest Losses Information System (APHLIS)[5] was also scanned for possible causes of losses in storage. The focus of the literature review was on farm storage losses for cereals and pulses in Sub-Saharan Africa. In total, forty journal articles were reviewed, covering eleven countries in the region and ten cereal crops, with maize being the most frequently assessed with 16 out of the 40 articles. Most articles used survey instruments to measure storage losses and their causing factors, while about seven articles were based on experimental designs that delved into the impact of storage infrastructure, packaging materials and the duration of storage on the level of grain losses.

General level of storage losses:

Grain stored on the farm can show a wide range of percentage loss levels, from less than 1% to over 20% or even 50% grain losses (Kumar et al., 2017; De Groote et al., 2023). Differences are observed between measurement and assessment methods. Wide ranges and high percentage losses are most likely to be found in controlled experiment-based studies, where the type of storage technologies are tested in a scenario of applying no interventions over a period of 6 to 9 months. These experiments can result in considerable insect and mould infestations in inappropriate storage conditions, while these are very low in hermetic storage (Ngwenyama et al., 2020.; Ng'ang'a et al., 2016). Model-based estimates provided by APHILS are summarized in Hodges et al. (2014). These are based on literature review and with recent values in APHLIS being modelled-based estimates, resulting in storage losses between 2-5% (Hodges et al., 2014). For Mali, APHLIS estimates maize household-level storage losses at 2.5%, while for millet, these are 0.3%. Survey-based estimates of storage losses show similar ranges. Abdoulaye et al. (2016), presents results from survey-based storage losses for Ghana, Benin, Burkina Faso, Nigeria, Ethiopia, Uganda and Tanzania, with estimates varying from a low of 1.9% in Burkina Faso to a high of 6.9% in Tanzania for maize, for legumes from a low of 1.3% in Burkina Faso to a high of 7.3% in Tanzania. Surveys on storage losses using physical measurement methods in Ethiopia and Ghana show similar results, with 2.66% of maize losses and 1.97% wheat losses in storage in Ethiopia (FAO/ESS, 2023), and 1.0% millet storage losses and 4.1% maize storage losses in Ghana (FAO, 2017).

Pest infestation and related causing storage conditions:

Insect infestation in storage is identified as one of the main primary causes of storage losses in Sub-Saharan Africa. Abdoulaye et al. (2016) shows that insects cause approximately 80% of maize storage losses in Ghana and Tanzania, between 70-80% of storage losses in Benin. The larger grain borer (*Proste-thanus truncatus*) is an important pest of on-farm stored maize, causing significantly higher losses than the more usual pests. Abass et al. (2013), in their studies on post-harvest food losses in a maize-based farming system in Tanzania, mentioned larger grain borer, but also rats, termites, microbes and toxins as the main causes. Apart from the larger grain borer, other insects infesting the grain are usually the grain weevil (*Sitophilus granarius*) and lesser grain borer (*Rhizopertha dominica*). Ratnadass et al. (1990) conducted surveys to assess losses caused by insect pests to sorghum grain stored in Malian villages, with the major pest encountered in all of them being the lesser grain borer. Nukeneni (2010), in a literature

[4] AGRIS Repository: https://www.fao.org/agris/publications
[5] APHLIS: https://www.aphlis.net/en

review concludes that insects are responsible for the greatest storage losses in cereals and pulses in SSA.

Insect pest infestation is often related to storage facilities not being able to maintain air tightness required to eliminate insect pests in storage. A relevant indication of the storage condition and resulting grain losses is the temperature and moisture level inside and outside the granary, as highlighted by Kumar et al. (2017). They mention that moisture content and temperature are the most crucial factors affecting the storage life of the grains.

Rodent attacks: While insects pests are widely outlined as the main source of grain losses during storage, Abdoulaye et al. (2016), in their survey results, identified rodent attacks as second most relevant cause of maize storage losses. Especially in Ethiopia, farmers declare these to account for about 40% of storage grain losses, in Uganda for about 20-30%. Ognakossan et al. (2016) conducted a survey in 2014 to assess magnitudes of postharvest losses in on-farm maize storage systems in Kenya and the contribution of rodents to the losses. They concluded that rodents represented the second most important cause of storage losses after insects, with farmers declaring that 45% of storage grain losses are caused by rodents.

Storage conditions: Given the problem related to pest infestation and rodent attacks, a major emphasis has been put in researching the impact of storage conditions on the level of pests causing grain losses. These are the type of storage facility, the type of packaging material the grains are stored in, but also the moisture content of grains and outside agro-climate conditions.

In terms of the storage facility, in most subsistence agricultures in sub-Saharan Africa, grains are mostly stored in the traditional structures at the household, to store the produce for consumption (reserves) and for seeds conservation (Adetunji, 2007). Usually, these storage structures are made of locally available materials (grass, wood, mud etc.). According to Abdoulaye et al. (2016), about 70-80% of farmers in the surveyed countries use either traditional storage or woven bags to store their grains. The authors further conclude that these storage conditions play a significant role in determining loss levels, impacting the level of grains having pest infestation or mould attacks.

Consequently, in the recent decade, improved storage infrastructure, but especially packaging innovations have been promoted by governments and development partners in many countries. Relative to packaging, losses differ between grains stored in packaged material and those stored lose on the ground (Kumar et al., 2017; Abass et al., 2013). Several studies outline how these technologies can help reduce losses by limiting the access of pests on the grains (Abass et al., 2014; Ngwenyama et al., 2020; Ng'ang'a et al., 2016). These studies, as already highlighted by Ratnadass et al. (1990), show that storage losses could be kept at very low percentage levels of below 1% over the storage period, if kept in adequate storage conditions. Therefore, grains stored in hermetic containers general have lower losses than those stored in ordinary packages (Kiaya, 2014). In this regard, packaging materials have been hypothesized to play an almost similar role for grain losses as the storage facilities and various researchers have looked at the impact of different packaging materials on loss reduction (Ogeudedji et al. (2018), Baributsa et al. (2020), Abdoulaye et al. (2016)). Ogeudedji et al. (2018), in their study conducted in Benin, observed that farmers who used plastic containers and storage bags had significantly lower losses than farmers who used cribs. Baributsa et al. (2020) report how maize farmers in Kenya were able to reduce pest infestation during storage by using hermetic bags.

Different storage practices, especially the use of storage chemicals, are also highlighted as factors that affect the storage losses (Manandhar et al., 2018; Kumar et al., 2017). Manandhar et al. (2018) evaluated the use of storage chemicals, with percentage of farmers responding positively from a low of 5% for legumes in Burkina Faso and 12% for maize in Uganda, to a high of 77% for maize in Ethiopia, indicating the extent to which the grains seem to be affected by pests during storage. The use of chemicals can trigger food safety problems if not used adequately, wherefore efforts have been made to use packaging materials rendering its use unnecessary.

Pre-storage conditions during harvesting and post-harvest:
Grain quality before storage is another relevant factor that can contribute to losses of grain during storage. Grain might be affected in post-harvest by inadequate drying, mechanical damage during harvesting and threshing, which can result in bruised areas on grains and points for infection and deterioration at storage (Kumar et al., 2017).

<u>Harvesting conditions:</u> In a study conducted in Pakistan, Sattar et al. (2015) analyse the effect of different harvesting methods on storage losses in wheat. Mechanized harvesting, if not done properly, caused grains breaking and therefore created an environment for entry of micro-organisms that increase losses later when the grain is stored. Also, the timing of harvest contributes significantly to losses experienced in storage. Abass et al. (2014) assessed the moisture level during harvesting and resulting losses during storage, with adverse effects if grains are harvested in rainy periods, affecting the moisture level of the grains and showing an increased susceptibility to pest infestation in storage.

<u>Post-harvest handling</u>: Similarly, poor post-harvest handling of grains can lead to an increase in losses during storage (Shee et al., 2019; Abass et al., 2014). Post-harvest activities may include drying, threshing, shelling, among others. Most smallholder farmers face challenges to properly handle grains after harvest and before storage. Baidhe et al. (2024) provide evidence for the link between drying and storage operations in the context of preserving grain quality. Shee et al. (2019) highlight lower losses in maize and sweet potato value chains in Uganda through training in proper post-harvest management. Adding on these, the weather conditions during post-harvest operations have also been identified as an explanatory variable for storage losses (Kumar et al., 2017)). APHLIS for instance recommends using rainfall data.

Duration of the storage period:
The period that the grains stay in storage also affects the loss levels and the probability of stored grain to be directly affected by bio-deterioration and pest infestation. Therefore, to understand and measure storage losses, the length of storage needs to be considered. The storage period can be short, with grains being stored for some weeks, while others reach a duration of more than 6-9 months. How long the grain is kept in storage depends highly on the use of the grains (for selling or for own consumption), the seasonality of grain production and availability, and the possibilities to mitigate the risk of grain losses. It has also been highlighted that losses can increase considerably when grains are stored for more than three to five months, especially losses caused by insect infestation, rodents, or deterioration and rotting that increases throughout the storage period (De Groote et al. (2013), Ngwenyama et al. (2020)).

Abdoulaye et al. (2016) indicate that some smallholder farmers in sub-Saharan Africa choose to sell their grains soon after harvest as a way to reduce losses. When farmers lack the means of investing in pest control mechanisms in storage, they resort to dispose the grains and sell them often at a lower price. Farmers are motivated to store their grains for longer periods if they have improved technologies that mitigate the impact of losses (Kadjo et al., 2018).

Hence, multiple factors impact on-farm grain storage losses, influenced by the storage conditions, like the storage infrastructure and packaging characteristics, but also the grain quality during storage influenced by the harvest and post-harvest management, and the external environment, like weather and general agroclimatic conditions. These can be further linked to the socio-economic characteristics of the farmers (e.g., age, sex, education level, experience, etc.), the political context (e.g., extension services, access to storage technologies, etc.), and market-related and consumption aspects (grain quality demanded in the markets, price and demand fluctuations, etc).

Table 1 aims at summarizing the main indicators and variables screened in the literature review.

Table 1: List of Explanatory Variables

Explanatory variable	Possible Responses	Sources
Storage facilities	1. Modern (Diffused light storage, metal silos…) 2. Traditional (granaries, mud huts, etc.)	Abass et al., 2014; Abdoulaye et al., 2016
Packaging type	1. PICS 2. Polypropeline sacks 3. Woven baskets 4. Synthetic bags 5. Plastic drums 6. Others (gourds, bottles,	Ogeudedji et al. (2018), Baributsa and Njoroge (2020), Abdoulaye et al., 2016.
Storage duration	1. Below 3 months 2. Between 3 to 6 months 3. Above 6 months	Kadjo et al., 2018; Abdoulaye et al. (2016)
Agro-Ecological Zones	Types of zones	
Knowledge/ Skills on Post Harvest Handling	1. Yes 2. No	Shee et al. (2019); Abass et al., (2019); Boxall 1998
Mode of harvesting	1. Mechanized 2. Manual	Sattar et al., (2015); Abass et al., (2019)
Time of harvesting	1. Early harvesting 2. Late harvesting	Abass et al., (2019)
Weather conditions at harvesting	1. Dry season 2. Rainy season	Abass et al., (2019)
Mode of threshing	1. Mechanized 2. Manual	Sattar et al., (2015)
Drying facilities	1. Modern driers	

	2. Raised racks 3. Polythene sheets 4. Floor drying 5. Others (drying in the farm, by the fireplace)	
Condition of commodity at storage	1. Shelled 2. Not Shelled	
Use of Pest controls e,g rodent traps, pesticides, insecticides Pest Incidence	1. High 2. Medium 3. Low	
Crop Variety	1. Traditional Variety 2. Hybrid/Improved varieties	Kadjo et al., 2018

2.3 Storage loss measurement and assessment methods

As highlighted in this chapter, losses during storage can occur in various moments and situations. In order to capture and quantify these, storage-specific measurement and assessment methods were developed. In general, three main assessment methods of grain storage losses can be highlighted, focusing on those that can be potentially used in larger sample surveys.

Declaration based:

Declaration-based storage loss assessment is less resource-intensive and more easily integrated into national farm surveys. However, it is prone to measurement errors and biases, as the accuracy of estimates depends on farmers' ability to observe, recall, and quantify their storage losses. As highlighted in GSARS (2018), storage losses might be too difficult to be declared by the farmers. One challenge lies in the calculation of percentage storage losses, given that storage loss reduction aims at reducing the loss percentage as a measure of efficiency. This means that precision is needed in estimating the storage loss quantities, but also in the corresponding total amount stored during the storage period. Given that grain is repeatedly removed and added to the storage, sometimes also from other harvests or bought from other farmers, the total amount stored and corresponding grain losses might be difficult to properly recall and aggregate over the whole storage period, with errors in either of these impacting results.

Grain losses due to insect attacks can be particularly difficult for farmers to observe and quantify. As previously mentioned, insects consume parts of the grain, reducing their weight. Estimating how many grains were affected and the extent of weight loss is often too complex to assess through visual observation alone.

Similar challenges arise with losses caused by microorganisms, which degrade grain quality and weight over time. In some cases, farmers may detect infestations (mould), discard the affected grain, and recall the incident. However, such occurrences are often sporadic and distributed over the entire storage period, making them hard to remember accurately. Losses due to rodents or birds are even less likely to be observed, as these animals often consume or remove grain without leaving clear traces.

In contrast, spillage losses are generally easier for farmers to observe, as they occur during handling—such as loading or unloading grain from storage. Nevertheless, spillage tends to be a less significant source of loss compared to other factors, except in systems where grain is stored in silos and handled mechanically. In such cases, spillage during transfer or grain left behind in equipment may become more relevant, although overall storage conditions are more controlled.

While in theory it might be possible to measure all these types of losses, doing so accurately through farmer recall remains highly challenging.

Apart from the difficulties to observe and quantity the grain losses, another limiting factor for farmers to recall the losses is the prolonged storage period, as well as the frequency of produce removals from or additions into storage. One strategy to mitigate these problems in the recall capacity of the farmers is to consider multiple visits during the storage period, with at least one visit immediately after harvest and another visit towards the end of the storage period between three to nine months after the first visit (Wollburg, 2021). In some situations, when farmers tend to constantly withdraw and add grain to the stock deposit during a prolonged storage period of more than four to six months, more than two visits might be required. GSARS (2018) recommends that data losses occurring during storage at the level of farm households is collected periodically, for example every month, for a one-year period.

If no additional field visits can be organized, the timing of asking for storage losses needs to be carefully considered. If the farm survey is conducted close to the harvesting period, storage losses will be low and result underestimated. Ideally, storage losses are to be declared at the end of the storage period. Given that this might not be feasible in a national farm survey, especially when storage periods differ between grain crops, regions and type of farmers, losses can only be declared for the former harvest and storage cycle, with probably further adverse implications for farmers to recall these.

Physical measurements:

Given the above-mentioned challenges for farmers to observe and quantify their grain losses during storage, research efforts have been put in place to develop physical measurement methods for grain storage losses. The summaries and descriptions of physical grain loss measurement methods provided by Boxall in 1986, Harris and Lindblad in 1978 and Compton in 1999 are still among the main references in this regard and fed into the GSARS guidelines (2018).

Physical measurement methods for grain storage losses aim to quantify weight loss, primarily resulting from damage caused by insects and microorganisms. These methods typically involve extracting grain samples from storage to assess the extent of damage - such as grains showing signs of insect infestation or fungal contamination. The observed damage is then converted into an estimate of equivalent weight loss.

The GSARS guidelines on post-harvest loss measurement for cereals (2018) summarize several methods following this general procedure, including: the Standard Volume/Weight Method (SVM), the

Conventional Count and Weight (or Gravimetric) Method, the Modified Count and Weight Method, the Thousand Grain Mass Method (TGM), and the Converted Percentage Damage Method. These methods mainly differ in terms of sampling techniques and how they calculate the equivalent weight loss of damaged versus undamaged grains.

Except for the SVM, most methods also include moisture content measurement, thus excluding weight losses due solely to water loss (rather than pest or microbial damage) from the final estimates. Additionally, these analyses typically exclude extraneous materials that may be present due to field or post-harvest handling.

Losses due to rodents and birds, as mentioned, can also contribute to on-farm losses of small-scale farmers. Unlike the case of insect pests or micro-organisms, there are no widely used assessment method specifically designed to measure losses from rodents and birds. Boxall (1986) differentiates its measurement whether the grain is stored on the cob or head or was threshed. With on cobs and heads, tested approaches allow its quantification with grain sampled. Threshed grain, on the contrary, would require comparing weights of grain stored and removed. Nevertheless, as highlighted by Kebe (2016), these can be challenging within farm-level studies because of the difficulty of monitoring all grain movements in and out of farm storage.

Storage loss estimates based on physical measurement methods—designed primarily to capture insect and microbial damage—tend to underestimate total losses because they do not account for grain lost to rodents and birds, spillage during handling, or infected grain that farmers may have removed and discarded. As such, a trade-off exists between the comprehensiveness of loss measurement and the cost of data collection. While physical methods provide reliable estimates for certain types of losses, they miss other significant loss pathways, such as those caused by rodents, birds, and handling errors.

The recommended physical measurement methods require at least two visits—and ideally three to four—to adequately cover the entire storage period, especially when storage lasts longer than three to four months. The first visit should take place at the beginning of the storage period to assess the initial quality of the grain. Subsequent visits should be spaced at least one month apart and timed to capture changes throughout the storage period. However, scheduling these visits can be challenging, as storage durations vary depending on the crop, region, and type of farmer. To ensure meaningful results, the visits should be timed to cover most of the storage duration for the majority of the grain stored.

Visual scales:

Most of the physical measurement techniques presented above involve collecting grain samples from the farmers, sending them to laboratories for analysis and later returning them. As highlighted in GSARS (2018), the implementation of laboratory analysis, as part of a national survey, means a considerable number of grain samples adding to the costs of the survey implementation. Also, the collection of the samples and their delivery to the laboratory represents an operational challenge, whereby grains need to be analysed timely to their collection to avoid the deterioration of the samples. These challenges triggered research on visual scales to be used as less resource-intensive method.

It has been shown that visual scales and standard charts offer a rapid and relatively accurate method for estimating storage losses directly in the field, without requiring laboratory analysis (Compton et al. 1998). Developed in the 1990s (Compton et al., 1991), these tools have been widely used to simplify data collection in loss assessments. Visual scales classify different levels of pest infestation in grains and provide corresponding visual representations. Through field experiments, each classification level is associated with an estimated percentage loss.

In practice, enumerators visually assess grain samples from storage and compare them to the standardized visual scale. Based on this comparison, they assign a classification level, which corresponds to an estimated percentage loss. The visual scales themselves are developed by subject matter specialists, while enumerators apply them in the field by matching samples to the reference images

Just like the other two approaches, several visits are necessary to fully capture the losses for the entire storage period.

Table 2: How different approaches capture losses based on their cause

	Insect storage losses	Rodent and bird storage losses	Deterioration storage losses	Handling losses
Declaration based	Difficult to observe, but experienced farmers might be able to declare	Difficult to observe, but experienced farmers might be able to declare	Partly declared, if farmers actively removed infested grains. Grain losses of infected grains cannot be observed and estimated.	Observable and farmers might be able to declare these
Physical measurements	Laboratory analysis	Not measured in laboratory analysis. Different methods used for grains on cobs/heads and threshed, therefore too complex to add.	Laboratory analysis	Requires observing the operation and weighing the spilled quantity out of the handled quantity. Therefore, often too complex to add.
Use of visual scales	Level of damage and corresponding losses are represented in the visual scales.	Eventually included when grain is stored on cobs or heads.	Level of damage and corresponding losses are represented in the visual scales.	Not possible to use visual scales

2.4 Resulting data needs on storage losses and driving factors

Based on the different types of grain storage losses, their underlying causes and drivers, and the available measurement methods, the data requirements for the grain storage loss survey module can be identified. To accurately estimate losses requires collecting a wide range of data, each essential for calculating the quantity of grain lost during storage.

Storage losses covering the storage period or parts of the storage period:

The central indicator to be collected is grain storage loss, defined as the quantity of grain exiting the food supply chain with no further use. When measured with the recommended physical methods, losses are estimated based on laboratory analysis. Key indicators assessed include grain moisture content, percentage of foreign material, and the weight and number of damaged versus undamaged grains.

To avoid underestimation of losses due to partial storage period coverage, declaration-based indicators—such as the planned storage duration—should also be collected. This section on storage losses should be completed at each visit, ensuring comprehensive tracking across the storage cycle.

Quantities stored and removed:

Understanding the quantities stored is essential for calculating loss percentages. In some cases, only part of the harvested crop is stored, meaning that using the total harvested quantity as a denominator can underestimate storage losses.

Stored quantities are collected through farmer declarations, along with information on grain added or removed between harvest and the first visit, and between subsequent visits. This information helps contextualize laboratory findings and estimate losses relative to stored volumes. Typically, stored quantities decline over time, while loss percentages increase as storage duration extends. These data points should be collected at each visit.

Storage infrastructure:

Since storage infrastructure is a major driver of grain losses, it must be thoroughly documented. Storage systems can vary widely, even within a country, and may be hard to classify. Surveys should collect information on materials used for the roof, floor, and walls, presence of temperature or humidity control, age of the structure, and maintenance frequency.

Data can be collected via enumerator observation, farmer declaration, and, when feasible, photos using Computer-Assisted Personal Interviewing (CAPI) tools. Because farmers often use multiple storage facilities, the survey should aim to capture information for each structure.

Condition of the grain during storage:

Post-harvest handling affects the grain’s condition and potential for losses. Common on-farm processing includes shelling, threshing, de-husking, cleaning, and drying. The survey should document whether grains were processed before storage and whether they are stored loose or in containers (e.g., bags, baskets). Practices like storing grain directly on the ground or elevating it can also impact losses and should be included.

The harvest and post-harvest handling activities are also important to capture during the survey. This will include the post-harvest activities performed on the commodity before storage and the type of equipment and technology used. They usually cover threshing/winnowing, de-husking, and drying. In terms of drying, the duration of the drying might be also captured, as well as the most common practices for drying the grains. The amount of grain losses that were incurred in these activities might also provide an indication on its efficiency and the overall post-harvest handling characteristics of the farm.

Weather variables:

Weather conditions are key explanatory variables. While basic weather data can be collected directly from respondents (e.g., at harvest or storage), geospatial weather databases now allow for more precise measurement. To use such data, GPS coordinates should be collected during field visits. Key weather indicators include temperature, humidity, and rainfall—all of which influence storage loss risks.

General farm characteristics:

Lastly, general socio-economic and structural characteristics of the farm and household are critical for contextual analysis. These are typically covered in national farm surveys and may include:

- Farm size
- Livestock ownership
- Machinery ownership
- Household demographics (e.g., age, gender, education)
- Farming experience
- Access to credit
- Household income

3. Implementation and results from the in-depth storage loss survey in Mali

3.1 Survey design of the in-depth storage loss survey in Mali

Based on the conceptual framework and the resulting data needs, a survey for the in-depth storage loss survey module has been designed and field tested in Mali. The survey was conducted to measure storage losses on the farm and gather and test potential factors that determine these losses. These variables were then used for modelling storage losses.

The in-depth survey on storage losses in Mali was conducted from February to April 2023 and included three visits. The pilot exercise focused on maize and millet. Maize is the most relevant cereal in terms of production, with 3 387 thousand tons produced in 2023, followed by rice with 3 024 thousand tons and milled 1 943 thousand tons (FAOSTAT, 2024). Given that in Mali, maize and millet are often cultivated together in a mixed farming system, it has been opted to focus on these two cereals to achieve a larger sample for each crop. The survey targeted two regions, namely Ségou and Sikasso and four Districts (Bla, Baroueli, Sikasso and Koutiala). The pilot survey was designed as a complementary module to the Enquête Agricole de Conjuncture[6] (EAC) survey 2023 conducted by the Cellule de Planification et de Statistique (CPS) du Ministère de l'Agriculture du Mali.

3.1.1 Sampling design

In this research, the main objective is to integrate the storage loss module in an existing national agricultural household survey, the EAC survey 2022/2023. To allow for physical measurement methods, a sub-sampling strategy has been followed as recommended in the GSARS guidelines (2018), whereby the storage loss module is only applied to a sub-sample of the households in the national agriculture survey. The sample followed the selection procedures in concordance with the EAC survey 2022/2023 sampling design. Three visits were conducted, aiming at covering the grain storage period, and importantly all sub-sampled holdings were to be included in all three visits.

The sub-sampling of holdings for the pilot survey was supported by the Cellule de Planification et de Statistique (CPS). The list of the enumeration areas of the four study districts and their respective households surveyed at the EAC survey 2022/2023 was used to select the enumeration areas for the pilot survey, summarised as follows:

Table 3: Sample size in-depth storage loss survey

Regions	Districts	Enumeration Areas	Households
2	4	48	288

In total, 48 enumeration areas were selected for the survey in 4 districts with an average of 12 enumeration areas were sampled per district. In each selected enumeration area, 6 agricultural households growing at least one of the two crops mentioned were chosen using a systematic sampling

[6] Please refer to https://www.instat-mali.org/fr/publications/enquete-agricole-de-conjoncture-eac

approach. The choice of households was favouring households that reported in the EAC to have grown both crops, although not all of them might store the produce.

In the proposed survey structure, two additional layers of sub-sampling were required within each sampled household. At the farm level, it was observed that farmers often use more than one storage facility. To enable the physical measurement of storage losses, it was necessary to select a single storage facility for grain sampling. A random selection mechanism was therefore integrated into the computer-assisted personal interview (CAPI) system to randomly choose the storage facility. In cases where a single storage structure contained both surveyed crops, the CAPI application allowed the same facility to be selected for both maize and millet

Based on the storage facility sampled, the last sub-sampling is done by randomly extracting the grain sample from the storage for the laboratory analysis. The procedures for doing so have been subject to research, given that the selection of the grains from the storage can introduce relevant biases. The corresponding guidance, tools and training materials on how to take the grain sample from the storage can be consulted in the GSARS guidelines (2018).

Given that not all households harvested both millet and maize, their final respective sample were lower than 288 households. Moreover, a number of farmers did not store grains or stored them for a short period only. This has led to replacements in the sampled households. Also, some farmers refused to participate which caused further replacements. A major difficulty occurred with famers storing their produce in hermetic bags meant for commercialization. These farmers were reluctant to provide access and open the bags to extract the grain sample, because it would increase the risk of potential contamination and deterioration of the grain. In some cases, CPS collected grain samples from secondary storage facilities, where these farmers stored the grain for own consumption.

The sample size was established based on available financial resources for the research, while aiming at achieving a minimum number of observations that would allow to test the modelling approaches. In general, the estimation of the minimum sample size before the survey design is complex and subject to research. Memon et al. (2020), for instance, provides an overall overview on the challenges and rules of thumbs used for determining the appropriate sample sizes, with a complex model using numerous variables requiring a larger dataset than a simple model with few variables. For panel data, some methodological texts recommend a minimum threshold of at least 50 cross-sectional units observed over 4 time periods to ensure sufficient variability and reliable estimation in panel data models (Hsiao, 2014). Here, the researchers worked with an approximated ratio of 50 observations for each explanatory variable in the modelling. The possible limitations of the sample size for the modelling exercise were observed only after data collection, when testing the different modelling approaches and identifying problems of multicollinearity that obliged to reduce the number of explanatory variables to about six explanatory variables.

3.1.2 Data collection instrument

The data collection instrument has been designed based on the data needs highlighted in chapter 2. The instrument for the pilot survey in Mali was designed to be complementary to the indicators already collected by the EAC survey 2022/2023.

The questionnaire had a core module to be applied only during the first visit and containing two main sections: (i) household/farm characteristics; and (ii) storage facility list and characteristics.

Section 1: Household/farm characteristics

- Cultivated area and Harvested area – For the three last agricultural seasons,
- Socio-economic information – At household and household head levels,
- Crop production – Crop harvested, sold and stored for the agricultural season of reference,
- Production techniques – Harvest, drying, threshing, and storage methods used.

Section 2: Storage facility

The section collected data on all the storage facilities at the household disposal. Questions on the structural characteristics of the facility, such as the type of walls, roof and the maximum capacity, are included. Information of crops stored in the facility at the time of the interview are also recorded. In this section, the CAPI questionnaire ran a random selection of the storage facility to be considered for the physical measurement. It also requested to take pictures of the storage facility to make sure the same facility was used for physical measurement in all following visits. In visits 2 and 3, information was only captured on the selected storage facility.

Section 3: Physical measurement of storage losses

Additionally, a separate section covers the indicators collected for the physical measurement of storage losses and corresponding movements of the quantities of grains stored, removed and added. The section captures all the changes in the amount of crop stored in the sampled storage facility only. It includes questions on the following quantities:

- Crop stored
- Crop withdrawn for consumption
- Crop sold
- Crop withdrawn due to damages
- Pest infestations (yes/no)
- Quantities removed and discarded

This module must be implemented in all visits. For the second and third visit, a reduced version has been used, focusing on the grain sample, grain removals or additions in the storage, incidence of pests and quantities removed to be discarded. On the other hand, the survey team identified key indicators to be added to the module for the following visits that are required to better capture storage dynamics. These were questions about moving grain quantities between storage facilities on the farm, the confirmation of the actually stored grains, as well as questions on the planned storage period to plan the next field visits.

Additionally, a specific sub-module was elaborated to be filled out directly by those conducting the laboratory analysis, including proper identification of the grain samples. The indicators obtained from the laboratory analysis are:

- Humidity percentage of the grain
- Percentage of impurity/strange/foreign materials
- Quantity and weight of damaged and undamaged grains

This design establishes that the same farms are surveyed repeatedly over time to track percentage grain losses and the corresponding stored quantities. As mentioned, the first visit provides an initial baseline to establish the quality of the grain in storage (in terms of humidity, impurity and damages), which becomes a reference point for subsequent assessments.

3.1.3 Data collection methods

The data collection was carried out through face-to-face interviews, using a computer-assisted personal interview (CAPI) software Survey Solutions[7].

Over the interviews, the enumerators had access to the selected storage facilities to collect the grains for the laboratory analysis. For each visit, 500 gr of the available crop(s) were collected. The sampling procedures took place using a dedicated drill probe with compartmentalized pockets. At the end of the crop collection operations, the sampled crops were conserved in a sealed envelope and stored in a dry place not directly exposed to the sunlight while waiting for transfer to the laboratory within a short time or at the end of the data collection of each visit. Samples arriving at the laboratory from the field were placed in a sealed moisture-proof container and stored at room temperature of the laboratory when opened. Therefore, it required prompt attention upon arrival by laboratory personnel.

During the first visit, the dehulled grains were sent to the laboratory as they were, which slowed down the work of the analysts. In the second and third visit, the dehulled sampled grains were hulled in the field before sealing them in the envelope. The grains sent to the laboratory were analyzed and the results recorded on cards. These sheets were collected and transcribed by the CPS team.

3.2 Survey field operation of the in-depth storage loss survey in Mali

The CPS/SDR recruited the enumerators among those who participated in the 2022-2023 annual agricultural survey in the study districts to facilitate the identification of households. Four teams, one team per district, were formed and each team consisted of a team leader and three investigators. In

[7] https://mysurvey.solutions/en/

addition, field supervision was planned during each visit. In total, 12 investigators, 4 team leaders and 4 field supervisors were mobilised for each visit which lasted one week between February and April 2023.

Two training courses were organised: training for supervisors and training for enumerators. The training of supervisors, provided by the AFRISTAT team and the World Bank, made it possible to train supervisors and interviewer team leaders at the CPS/SDR level for four days. Once the training of trainers was completed, the CPS trainers proceeded with the training of the enumerators under the supervision of the team formed by AFRISTAT, FAO and the World Bank.

Data collection involved three visits to the field to gather crop samples for laboratory analysis. Planning these visits proved challenging, as limited prior information was available regarding the typical storage duration of maize and millet in the selected areas. The goal was to cover the full storage period; however, it was also important to maintain sufficient sample sizes across all three visits. This required avoiding cases where farmers had already depleted their stored grain by the time of the second or third visit. To mitigate this risk, a conservative approach was adopted in estimating storage duration, based on the assumption that most farmers would store their grains for approximately three to four months. The first visit was conducted from February 9th to 16th. The second visit took place from March the 12th to 19th, while the third visit started on April 8th and concluded on 15th.

Table 4 summarize the data collection results for the three visits. In visit 1, all 288 households were positively interviewed, while in visit 2 one household refused to be interviewed and in visit 3 three other households decided not to participate in the interviews. The household in visit 2 that refused to participate was dropped from the subsequent visit, while those refusing to participate in visit 3 were kept given that two subsequent observations were available. The following tables show the result of the sample achieved for each crop in the four districts.

Table 4: Total household sample Size by district and crop during the three visits

	Visit 1	Visit 2	Visit 3
District	***# Households interviewed***	***# Households interviewed***	***# Households interviewed***
Koutiala	72	72	71
Sikasso	72	71	71
Baroueli	72	72	70
Bla	72	72	72
Total HHs	**288**	**287**	**284**
% of total	**100.0%**	**99.65%**	**98.65%**

Table 5 shows the data collection results for all districts, summarizing the number of households visited in each visit and samples obtained by crops. Not all households surveyed cultivate both crops: out of the 288 farmers, 234 cultivated millet and 187 maize. For the second visit, apart from the one household that refused to participate, the number of crops sampled decreased due to the end of stock availability in eight households of Koutiala district (6 of millet and 2 of maize), in three households in Baroueli (maize) and in five households in Bla (5 stocks of maize and 3 of millet).

During third visit, apart from three additional households who refused to be interviewed, the total amount of crops sampled was reduced by three more farmers in millet and maize due to a further decrease in stock availability for the interviewed households. Overall, compared to the previous visit, six samples for both millet and maize were missing. However, no households were dropped from the survey given that at least two visits were concluded for these households, which is the minimum required to calculate incurred storage losses.

Table 5 presents the final observations by district, crop and visit. In total, 1 209 grain samples (525 corn samples and 684 millet samples) were collected and sent to the laboratory for analysis. The achievement compared to the forecast is overall 70%: 61% for maize and 79% for millet.

Table 5: Total number and distribution of grain samples by district and crop

	No of HHS			Number of grain samples								
				Maize			Millet			Total		
Districts	**Vis 1**	**Vis 2**	**Vis 3**	**Vis 1**	**Vis 2**	**Vis 3**	**Vis 1**	**Vis 2**	**Vis 3**	**Vis 1**	**Vis 2**	**Vis 3**
Koutiala	72	72	71	60	54	51	65	63	59	125	117	110
Sikasso	72	71	71	71	70	70	28	27	26	99	97	96
Baroueli	72	72	70	17	14	12	69	69	68	86	83	80
Bla	72	72	72	39	34	33	72	69	69	111	103	102
Sub-total	288	287	284	187	172	166	234	228	222	421	400	388
Total				**525**			**684**			**1290**		

3.3. Results from the in-depth storage loss survey module in Mali

3.3.1 Estimation of stored loss

As outlined in chapter 2, the physical measurement method employed here is designed to determine—through laboratory analysis—the quantity of grain affected by pest infestation and deterioration, along with the corresponding weight loss. This method has been operationalized into a calculation approach that also accounts for moisture loss and the removal of foreign materials. The approach is based on the Count and Weight Method (Harris & Lindblad, 1978). The procedure involves separating grain for each sample into undamaged and damaged portions, and measuring their weight difference with the following formula:

$$\%\,weight\ loss\ = \frac{((UNd) - (DNu))}{U(Nd+Nu)} * 100$$

Where U = weight of undamaged grain,

Nu = number of undamaged grains,

D = weight of damaged grains,

Nd = number of damaged grains.

Figure 2 presents the results for storage grain losses by each of the visits for millet and maize.

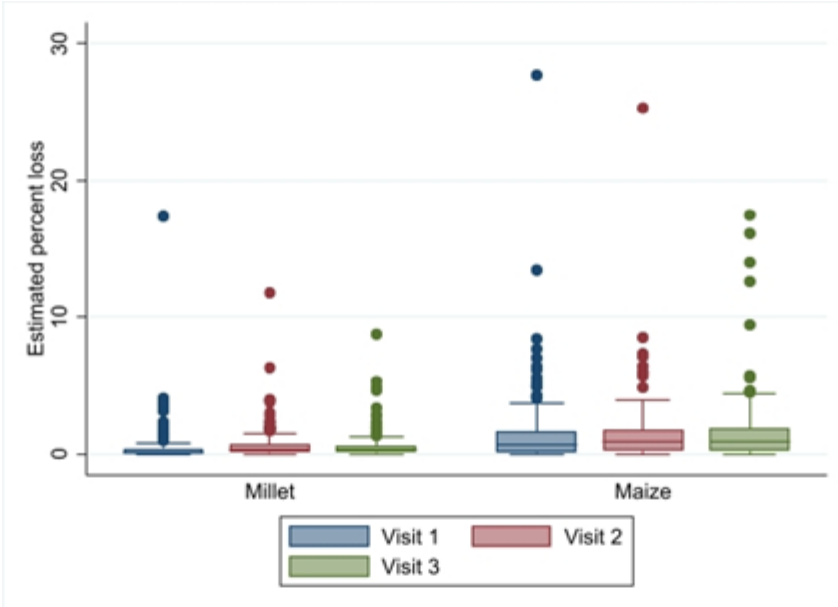

Figure 2: Estimated loss percentage by crop and visit

Grain losses are higher for maize compared to millet. In both cases, as expected, losses increase with the time in storage, in the case of millet from 0.4% in the first visit to 0.6% in the third visit; maize increased from 1.3% to 1.6%. Both crops show similar asymmetrical distributions by visit. Percentage losses result to be higher for more households with each visit, while extreme values are less prevalent by the third visit, probably due to consuming or selling the produce before incurring more losses. Mean, standard deviation and quartiles can be found in Table 6.

Table 6: Descriptive Statistics for Percentage Loss by Crop and Visit

Crop	Visit	Mean	Std. Dev.	Minimum	P 25	P 50	P 75	Maximum
Millet	1	0.4	1.3	0	0.02	0.07	0.4	17.4
	2	0.6	1.1	0	0.1	0.3	0.7	11.8
	3	0.6	0.9	0	0.1	0.3	0.6	8.8
Maize	1	1.3	2.6	0	0.1	0.6	1.6	27.7
	2	1.4	2.4	0	0.2	0.9	1.7	25.3
	3	1.6	2.6	0	0.2	0.9	1.9	17.5

3.3.2 Basic Descriptive Statistics

Table 7 to Table 12 provides descriptive statistics regarding the main characteristics of the interviewed households.

Table 7 shows average estimates of the cultivated and harvested area for both millet and maize for the investigated agricultural season as well as the two seasons prior to the data collection. Although the cultivated and harvested area for millet is generally twice the area of maize, on average the trend over the years is very similar for both crops.

In the case of millet for the 2022/23 season, on average, the harvested area stands at 3.70 hectares (91.17% of the cultivated area), with a noticeable disparity between Ségou (4.73 ha) and Sikasso (2.24 ha). In contrast, maize exhibits a lower average harvested area of 1.98 hectares (but a higher share of harvested over cultivated area, 95.7%), with Sikasso (2.75 ha) surpassing Ségou (0.92 ha).

The distribution of cultivated areas across different size categories provides insights into the scale of farming operations. Across all seasons and regions, most of the households' activity fell within the 1 to 3-hectare category, indicating the prevalence of small to medium-scale farming practices.

Regarding the harvested area, for millet, the main range is the 3-to-6 hectares (37.5% of the households) followed by the 1-to-3 hectares (36.6%), with significant differences between regions. In Sikasso, 54.5% of the households harvested an area between 1 and 3 hectares, while 35.7% harvested a bigger area. On the other side, in Ségou, most of the households harvested between 3 to 6 hectares of area (40.9%) or less than 3 hectares (26.8%). For maize, the trend is similar in both regions, with a prevalence of harvested areas in the 1-to-3 hectares size range. However, while in Sikasso almost 60% of the households fell within this category, in Ségou the prevalence of the harvest activities was carried on in plots smaller than 1 hectare (50.5% of the households).

Table 7: Average area cultivated and harvested

	Millet			Maize		
	Total	**Ségou**	**Sikasso**	**Total**	**Ségou**	**Sikasso**
Agricultural Season 2020-2021						
Cultivated area (avg, hectars)	**3.98**	**5.09**	**2.58**	**2.02**	**1.03**	**2.78**
Cultivated area range						
< 1 hectars	2.1%	0.7%	3.8%	22.6%	42.6%	7.1%
1 to 3 hectars	34.0%	15.6%	57.6%	56.9%	53.7%	59.3%
3 to 6 hectars	41.5%	49.6%	31.1%	15.3%	2.8%	25.0%
6 hectars or more	22.4%	34.1%	7.6%	5.2%	0.9%	8.6%
Harvested area (avg, hectars)	**3.78**	**4.83**	**2.45**	**1.97**	**1.00**	**2.71**
Harvested area range						
< 1 hectars	2.1%	0.7%	3.8%	23.6%	44.3%	7.9%
1 to 3 hectars	39.8%	21.5%	63.2%	55.7%	51.9%	58.6%
3 to 6 hectars	37.3%	46.7%	25.5%	15.9%	3.8%	25.0%
6 hectars or more	20.8%	31.1%	7.6%	4.9%	0.0%	8.6%

Agricultural Season 2021-2022						
Cultivated area (avg, hectars)	**3.92**	**5.15**	**1.33**	**1.99**	**0.97**	**2.76**
Cultivated area range						
< 1 hectars	2.6%	0.8%	5.0%	23.2%	45.7%	6.4%
1 to 3 hectars	36.3%	17.9%	61.0%	59.4%	51.4%	65.3%
3 to 6 hectars	40.6%	47.8%	31.0%	13.0%	2.9%	20.6%
6 hectars or more	20.5%	33.6%	3.0%	4.5%	0.0%	7.8%
Harvested area (avg, hectars)	**3.63**	**4.67**	**2.21**	**1.92**	**0.90**	**2.66**
Harvested area range						
< 1 hectars	3.0%	1.5%	5.1%	24.5%	47.5%	7.9%
1 to 3 hectars	42.1%	26.9%	62.6%	58.5%	50.5%	64.3%
3 to 6 hectars	37.8%	44.0%	29.3%	12.5%	2.0%	20.0%
6 hectars or more	17.2%	27.6%	3.0%	4.6%	0.0%	7.9%
Agricultural Season 2022-2023						
Cultivated area (avg, hectars)	**4.05**	**5.24**	**2.38**	**2.07**	**1.01**	**2.86**
Cultivated area range						
< 1 hectars	4.9%	2.8%	7.9%	23.3%	45.2%	7.1%
1 to 3 hectars	31.7%	16.2%	53.5%	56.7%	49.0%	62.4%
3 to 6 hectars	41.6%	45.8%	35.6%	13.9%	5.8%	19.9%
6 hectars or more	21.8%	35.2%	3.0%	6.1%	0.0%	10.6%
Harvested area (avg, hectars)	**3.70**	**4.73**	**2.24**	**1.98**	**0.92**	**2.75**
Harvested area range						
< 1 hectars	5.8%	2.8%	9.9%	27.4%	50.5%	10.7%
1 to 3 hectars	36.6%	23.9%	54.5%	53.5%	45.5%	59.3%
3 to 6 hectars	37.5%	40.9%	32.7%	13.3%	4.0%	20.0%
6 hectars or more	20.2%	32.4%	3.0%	5.8%	0.0%	10.0%

Table 8 provides data on production, sales, and stocks stored. The total average production for interviewed households is higher for maize than for millet, although households that cultivated maize crop are fewer in number than those cultivating millet. In general, the main purpose of the cultivation is self-consumption. Overall, only 20% of millet production has been intended for sale, while the share is slightly higher for maize (29.7%). No significant differences have been found when analysing regional trends.

In terms of production quantity, most of the interviewed households produced, on average, between 1000 and 4000 kgs for both crops (53.7% for millet and 43.6% for maize). Differences have been found at the regional level: households in the Ségou region have a clear preference toward cultivation of millet (average production is more than the double of production in Sikasso), maize crop is the main one for households residing in Sikasso when compared to millet

Concerning sales, more than one out of three households decided to sell less than 500 kgs. Despite larger sales (on average) of maize crop, the share of households that sold more than 2000 kgs of crop is higher for millet (5.5%, all due to sales in the Ségou region), than for maize crop (4%).

Table 8: Data on production, sales and stock volumes and distribution by ranges

Production, sales and stocks for the 2022-23 season	Millet			Maize		
	Total	Ségou	Sikasso	Total	Ségou	Sikasso
Production (avg, *'000* kg)	**2.66**	**3.43**	**1.57**	**3.20**	**1.23**	**4.58**
Production range						
<500 kg	9.9%	1.4%	22.0%	15.7%	28.9%	6.5%
500 to 1000 kg	14.1%	9.9%	20.0%	16.6%	25.8%	10.1%
1000 to 2000 kg	22.3%	17.6%	29.0%	25.1%	27.8%	23.2%
2000 to 4000 kg	31.4%	38.0%	22.0%	17.5%	13.4%	20.3%
4000 kg or more	22.3%	33.1%	7.0%	25.1%	4.1%	39.9%
Sales (avg, '000 kg)	**0.52**	**0.63**	**0.36**	**0.95**	**0.38**	**1.34**
Sales range						
<500 kg	67.0%	61.1%	75.7%	73.7%	75.0%	72.9%
500 to 1000 kg	17.6%	20.4%	13.5%	18.2%	22.5%	15.3%
1000 to 2000 kg	9.9%	9.3%	10.8%	4.0%	0.0%	6.8%
2000 to 4000 kg	3.3%	5.6%	0.0%	3.0%	2.5%	3.4%
4000 kg or more	2.2%	3.7%	0.0%	1.0%	0.0%	1.7%
Storage (avg, '000 kg)	**2.13**	**2.72**	**1.26**	**2.57**	**0.99**	**3.48**
Storage range						
<500 kg	16.0%	5.7%	30.9%	22.9%	46.7%	9.2%
500 to 1000 kg	18.5%	14.2%	24.7%	17.6%	20.0%	16.2%
1000 to 2000 kg	21.9%	19.9%	24.7%	23.4%	24.0%	23.1%
2000 to 4000 kg	29.8%	41.1%	13.4%	16.6%	6.7%	22.3%
4000 kg or more	13.9%	19.2%	6.2%	19.5%	2.7%	29.2%

Table 9 provides data on the threshing process for millet and maize, including the duration of threshing and the method used. The average threshing duration is 1.33 days overall for millet and 1.40 days for maize. Differences at a regional level are very limited. For millet, it takes 1.26 days in Ségou, while in Sikasso it takes 1.43 days. The trend is the opposite when analysing maize crop: in Ségou duration was on average 1.52 days, whereas Sikasso has a shorter duration at 1.28 days. With respect to the threshing methods, almost 4 out of 5 households used a modern (with the help of machinery) threshing method for both crops. Slight variations exist between regions, with Ségou generally showing a higher incidence of traditional/manual methods (24.7% for millet and 27.8% for maize) compared to Sikasso (14.6% and 15.2%, respectively).

Drying duration is generally higher for maize when compared to millet (13.6 against 9.4 days, on average). The majority of drying periods fall within the 6 to 10 days range for both Millet and Maize, even if some variations between regions and crops persist. Few households in the sample did not implement any drying representing 2.1% for millet and 1.64 for maize.

For millet, almost all the interviewed households used floor drying to dry the harvested crop (94.2%), with almost nobody adopting raised grills and polyethylene sheets drying methods (0.8% for both) and very few preferring other drying methods. The trend is confirmed for maize, although a higher share of households dried the harvested crop using the polyethylene sheets method (18.7%) and other drying methods. In addition, for maize, differences between regions have been found, with a clear predominance of floor drying technique in Ségou (87.6%) compared to Sikasso (only 48.6% of the households).

Table 9: Threshing and drying characteristics by crop

	Millet			Maize		
	Total	Ségou	Sikasso	Total	Ségou	Sikasso
Threshing						
Threshing duration (avg, days)	1.33	1.26	1.43	1.40	1.52	1.28
Threshing method						
Traditional/manual	20.6%	24.7%	14.6%	21.7%	27.8%	15.2%
Modern machine	79.4%	75.4%	85.4%	78.3%	72.2%	84.8%
Drying						
Drying duration (avg, days)	9.44	9.71	9.04	13.63	13.91	13.43
Duration range						
< 6 days	26.0%	26.8%	25.0%	13.6%	9.3%	16.7%
6 to 10 days	46.3%	45.8%	47.0%	35.3%	34.0%	36.2%
11 to 15 days	13.6%	10.6%	18.0%	23.8%	27.8%	21.0%
16 days or more	14.1%	16.9%	10.0%	27.2%	28.9%	26.1%
Drying methods						
No drying	2.1%	1.4%	3.0%	2.1%	3.1%	1.5%
Raised grills	0.8%	0.0%	2.0%	6.0%	5.2%	6.5%
Polyethylene sheets	0.8%	0.0%	2.0%	18.7%	4.1%	29.0%
Floor drying	94.2%	97.9%	89.0%	64.7%	87.6%	48.6%
Other	3.3%	0.7%	7.0%	8.5%	0.0%	14.5%

Table 10 provides data on the distribution of storage methods utilized for millet and maize, categorized by different types of storage bags. It presents percentages representing the prevalence of each storage method in total storage for both crops, specifically distinguishing between the Ségou and Sikasso regions.

While polypropylene bags dominate as the primary storage method, other options such as jute bags, and hermetic storage bags are also employed in both regions. However, the most common storage method is the one loose on the ground, adopted by almost one out of two households for both millet and maize.

Table 10: Storage methods used for storing maize and millet

Storage bag	Millet			Maize		
	Total	Ségou	Sikasso	Total	Ségou	Sikasso
Jute bags	2.5%	0.7%	5.2%	2.7%	0.0%	3.9%
Polypropylene (PP) bags	30.8%	22.9%	32.3%	41.2%	39.7%	41.9%
Hermetic storage bags	3.4%	5.0%	1.0%	2.7%	8.6%	0.0%
Hermetic metal silo	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Hermetic plastic silo	0.4%	0.7%	0.0%	0.0%	0.0%	0.0%
Loose on the ground	46.4%	42.9%	51,55%	46.5%	29.3%	54.3%
Other	16.5%	27.9%	0.0%	7.0%	22.4%	0.0%

Table 11 shows data on the stored quantities at visit 1 and the changes in the total amount declared by respondents during the three visits.

On average, the stored quantity at the first visit was higher for maize (more than 2000 kgs) than for millet (1601.84 kgs). However, with maize being a major staple crop in Mali, the average quantity consumed by the households between harvest and the first visit is three times higher than millet, which is shown by the results in Table 14, indicating that a higher proportion of households sold smaller amounts of corn than of millet. In addition, this data reflects the difference in the harvest calendar for the two crops, with maize being the first crop to be harvested and, consequently, consumed (the maize harvest took place between September and October), while on average millet harvesting ended one month later (mid-November to December). This might mean that millet was still at the beginning of the storage period, with a probability of grain losses to increase in the later months, while for maize, the visits were conducted in later moment of the storage period. However, data at the regional level shows that almost all the quantity damaged is coming from the region of Sikasso (363 kgs per household) while households in Ségou discarded only 7.50 kgs each.

Moreover, data regarding the period between the harvest and the first visit shows that crop quantities given out for other reasons (such as gift or in-kind payments among others) have been on average very consistent (more than 250 kgs per household).

For the stock variation between the three visits, the trend is generally similar for both crops. The level of crop added coming from household cultivations is very low in both periods between visits 1 and 2, and between visits 2 and 3, while for both visits 2 and 3, the main entry for millet and maize is the quantity consumed. However, for the case of maize, the period between visits 2 and 3 saw a rise in the quantity sold, mainly due to sales in Sikasso (168 kgs per household).

Table 11: Stored quantities by visit and changes in stock between visits

Stored quantities and changes between visits	Millet			Maize		
	Both	Ségou	Sikasso	Both	Ségou	Sikasso
Harvested quantity (avg, 1000 kg)	**2.66**	**3.43**	**1.57**	**3.20**	**1.23**	**4.58**
General harvest period		**Oct/Nov**	**Sept**		**Nov/Dec**	**Sept**
Variation between harvest and V1 (avg, kg)						
Consumed	520.33	566.88	479.00	1,738.77	616.28	2,260.23
Sold	526.81	573.19	412.10	569.26	432.24	744.22
Damaged grains	48.49	54.62	45.77	350.04	7.50	363.74
Other	273.12	317.00	193.92	281.05	210.01	315.50
Stored quantity at visit 1 (avg, kg) -	**1,601.84**	**2,004.26**	**1,021.03**	**2,057.89**	**676.54**	**2,678.97**
Variation between V1 and V2 (avg, kg)						
Added to storage	12.13	18.87	2.20	22.57	0.00	32.87
Consumed	214.42	236.53	181.99	305.89	108.76	395.92
Sold	43.98	48.21	37.75	26.83	56.44	13.31
Damaged grains	1.79	0.34	3.91	17.07	0.34	24.71
Other	19.66	29.60	5.00	10.21	19.45	5.98
Stored quantity at visit 2 (avg, kg)*	**1,334.12**	**1,708.45**	**794.58**	**1,720.46**	**491.55**	**2,271.92**
Variation between V2 and V3 (avg, kg)						
Added to storage	1.75	2.93	0.00	2.70	0.00	3.91
Consumed	175.77	193.91	148.47	161.83	50.38	211.76
Sold	25.26	31.09	16.49	122.98	21.55	168.42
Damaged grains	0.37	0.61	0.00	0.07	0.23	0.00
Other	10.69	16.68	1.69	0.85	2.58	0.08
Stored quantity at visit 3 (avg, kg)*	**1,123.78**	**1,469.09**	**627.93**	**1,437.43**	**416.81**	**1,895.57**

** calculated based on the average quantities added and removed*

3.4 Lessons learnt from the in-depth storage loss survey

Several lessons can be derived from the exercise on designing and implementing an in-depth storage loss survey module. First, it allows to collect key information about the existing storage infrastructure on the farm, the number of storage facilities used, the type of storage and package technology, as well as the quantities stored and corresponding in- and outflows throughout the observed storage period. These are relevant indicators for informing any intervention to improve on-farm storage, targeted by policies to reduce food losses, but also to improve food safety related to aflatoxin prevalence, or understand food security and income patterns through the consumption and selling patterns observed throughout the storage period. Inserting this specific module into national agriculture surveys could fill important data gaps, given that detailed information on storage characteristics is mostly scattered and only available for few observations through case studies or controlled experiments.

In terms of measuring grain storage losses, the results obtained are aligned with the literature, where maize has usually higher percentage losses in storage compared to millet (APHLIS). Given Mali's agroclimatic conditions, storage grain losses are generally observed to be lower compared to other more tropical countries in the region. APHLIS estimated 2.5% of storage losses for maize, and 0.4% of millet (APHLIS). The obtained storage loss estimates also show relatively low standard errors, especially observed to declaration-based estimates of storage losses (FAO and ESS, 2023). Nevertheless, it is important to acknowledge that physical measurements based on laboratory analysis, while these are precise in measuring losses from insect attracts and moulds, do not cover all quantity losses (i.e. rodent attacks damaging packaging and contaminating grains, removed and discarded quantities during storage due to pest and mould infestation, discarded and rejected grains after storage). Farmers' declarations on the quantities removed and discarded were collected, but since these happen in sporadic manner, too few observations were obtained to be used as part of the storage loss estimation and the modelling exercise. Further research might be needed in Mali to assess whether other main loss points in grain storage considerably add to the loss levels and should therefore be included.

In terms of the results obtained on grain storage percentage losses, the percentage storage losses in the three visits did not increase in the pace as it is observed in most experimental studies on storage losses. This is due to some extent to the short interval between visits, but also because units with high observed percentage losses in the first visit were removed at the second and third visit, indicating that farmers incurring high percentage losses might have sold or consumed the produce to avoid further losses. These decision-making dynamics can result in an overall downward trend of average percentage storage losses throughout the storage period, while percentage storage losses of each farmer must either stay constant or increase with time the grain is stored (damaged grain cannot be reverted).

While most of the indicators showed met expected results in terms of the level of grain losses, some were counterintuitive, as for lower storage grain losses of grains stored loose on the ground compared to packed grains. With most farmers reporting to store lose on the ground, further specifications might be needed to understand the factors protecting the grain in these storages compared to those stored in jute bags and hermetic bags.

A major challenge of the study was related to the conduct of laboratory analysis and to the various visits. Choosing the timing of storage losses measurement and visits added complexity. Ideally visits should span across the whole storage duration, with a first visit after harvesting and the last visit happening one month before ending grain storage. This is a challenge when considering that crops are harvested at different times by regions and farmers, and so are the planned storage periods. It is therefore important to obtain specific information on the duration of storage and related dynamics of grain consumption and commercialization beforehand, for planning the storage loss survey. This has been a limitation to the exercise in Mali, that started with the first visit some months after harvesting and was closed with the last visit when almost 50% of the harvested grain was still in storage, and three to four months storage were still ahead.

Another operational complexity was the number of grain samples to be analysed in a laboratory in a relatively short time. The method therefore requires an agreement of the government institutions with regional laboratories, assessing their installed capacities to handle the amount of collected grain samples. Also, the samples need to be processed within a critical period after data collection, which proved challenging due to the operational arrangements to bring the samples from the rural areas to the laboratories. As alternative solution, staff from the national statistics office can be trained or hired to conduct the analysis in the field. Looking at future exercises, visual scales could be calibrated for direct use by the enumerators in the field, representing a much cost-efficient method on a large-scale national farm survey.

4. Storage loss models and their areas of application

4.1 Data preparation

The proposed models in this study are intended to reach two main goals at the same time. First the establishment of useful models to identify main drivers for storage losses based on theoretical knowledge, focusing on drivers that could be relatively easy to measure and to be integrated in household or farm surveys. Second, the assessment of model prediction capabilities to estimate the percentage of storage losses. This prediction capabilities, can then be used to estimate grain storage losses indirectly, using the set of explanatory variables collected in the national household survey.

The establishment of on-farm loss models for stored millet and maize in Mali, based on a 3 visits data sample, requires the integration of several data sets, data cleaning and an exploratory analysis. Data sets were obtained from three different sources:

- Data from the implemented in-depth storage loss survey and the corresponding laboratory results
- Indicators collected during EAC survey 2022/2023 (Enquête Agricole de Conjoncture Intégrée aux Conditions de Vie des Ménages 2022/2023), given that the in-depth storage loss survey used a sub-sample of the main survey
- Weather variables from Copernicus Climate Change Service using the GPS information collected on the farm

Data from the in-depth storage loss survey in Mali covered the most ad-hoc and storage-specific set of indicators, such as the structure of storage facilities, loss data obtained from the questionnaires applied during the three visits, harvesting and pre-storing conditions, as well as and socio-economic, agricultural and regional factors. This dataset was cleaned and structured for data analysis. Laboratory data form the three visits, had to be processed and integrated to the dataset, with the obtained grain storage losses in percentages. Adding on these, a set of selected variables from the EAC survey 2022/2023was provided by CPS and integrated to the data set. Weather variables were downloaded from the publicly available Copernicus Climate Change Service, processed and jointed to an integrated database.

Variables from Mali in-depth storage loss survey 2023

Based on the previous literature review, the selection of potential loss determinants was established, where variables were arranged by main topics. These potential predictor variables collected in the in-depth storage loss survey in Mali are presented in Table 12.

Table 12: Selection of potential explanatory variables for grain storage losses

Questionnaire predictors by topic	Variable	Content
Module HARVESTED AREA (First visit)		
3.01c. How much area was CULTIVATED for the crop during the 2022/23 agricultural season? (HECTARES)	s03q01c	Cultivated area in hectares
3.02c. How much area was HARVESTED for the crop during the 2022/23 agricultural season? (HECTARES)	s03q02c	Harvested area in hectares

Module SOCIO-ECONOMIC (First visit)		
3.04. Age of head of agricultural household (YEARS)	s03q04	Age in years
3.07. How many years of experience does the head of household have in agriculture?	s03q07	Experience in years
3.08. What is the annual income level of the head of the agricultural household?	s03q08	Income in FCFA currency
3.09. Does the head of household generally have access to credit?	s03q09	Access to credit (yes=1, no=0)
3.10. What is the number of permanent agricultural workers in the household?	s03q10	Number of permanent workers
3.11. What is the number of temporary agricultural workers in the household?	s03q11	Number of temporary workers
3.13a. Have you received training on practical techniques to reduce post-harvest losses?	s03q13a	Received training (yes=1, no=0)
3.13b. Do you think you have properly assimilated and used the techniques learned during the training on post-harvest losses received?	s03q13b	Properly training assimilation (yes=1, no=0)
3.14. How much is harvested during this agricultural season? (kg)	recolte_kg	Grains harvested in kilograms
3.16. How much is stored during this agricultural season? (kg)	stockee_kg	Grains stored in kilograms
Module USED TECHNIQUES (First visit)		
3.21. Did you start harvesting early (in advance) compared to past years?	s03q21	Start harvesting early (yes=1, no=0)
3.22. Did you start harvesting later than in past years?	s03q22	Start harvesting later (yes=1, no=0)
3.23. Has there been enough rain to negatively impact harvesting activities?	s03q23	excessive rain impact (yes=1, no=0)
3.20. How many days did the harvest last?	s03q20	Harvest duration in days
3.27. What method did you use to thresh?	s03q27	Treshing method (1=Manual, 2=Modern machine)
3.30 How many days did it take for the grain to dry?	s03q30	Drying duration in days
3.31. What method are you sometimes using to check the crop? 00 No drying	s03q31__0	Method used to check the crop: No drying (yes=1, no=0)
3.31. What method are you sometimes using to check the crop? 01 Raised racks	s03q31__1	Method used to check the crop: Raised racks (yes=1, no=0)
3.31. What method are you sometimes using to check the crop? 02 Polyethylene bags	s03q31__2	Method used to check the crop: Polyethylene bags (yes=1, no=0)
3.31. What method are you sometimes using to check the crop? 03 Floor drying	s03q31__3	Method used to check the crop: Floor drying (yes=1, no=0)
Module STORAGE FACILITY STRUCTURE (First visit)		
4.02. What type of walls has the structure?	s04q02	Walls of the structure (recoded to: 1=Bench bricks, 2=Woven basket, 3=Mud, 4 Other)

4.03. What type of floor/ground has the structure?	s04q03	Floor/ground of the structure (recoded to: 1=Concrete, 2= Earth, 3=Woven basket, 4=Wood, 5=Other)
4.04. What is the height of the platform relative to the ground level of the structure?	s04q04	Height of the platform (recoded to: 0 =ground level, 1= 0.5 meters, 2=1+ meters)
4.05. How old is the structure?	s04q05	age of the structure in years
4.06. What was the cost of the structure? (FCFA)	s04q06	Structure cost in FCFA
4.07a. What is the maximum storage capacity of the structure? (QUANTITY)	s04q07a	Structure capacity in kg
4.12c. How much culture is currently stored in structure? Kg equivalence	s04_kg	Stored amount in kg
Module LOSS ESTIMATION (First, second & third visit)		
5.04. What type of container did you use for the culture stored in this structure?	s05q04	Typr of container (1=Jute bags, 2=Polypropylene (PP) bags, 3=Hermatic storage, 4=Loose on the ground, 5=In bulk, 6=On cobs, 7=Bank, 8=other)
5.05d. What is the quantity currently stored in this structure? Kg equivalence	s05q05d_1	Stored quantity in kg
5.07d. What is the amount removed for consumption in this structure? Kg equivalence	s05q07d_1	Consumed quantity in kg
5.08d. What is the quantity withdrawn for sale in this structure? Kg equivalence	s05q08d_1	Sold quantity in kg
5.14d. What is the amount removed due to spoiled grain in this structure? Kg equivalence	s05q14d_1	Damaged quantity in kg
5.15d. What is the amount removed for any other reason in this structure? Kg equivalence	s05q15d_1	Retired quantity other causes in kg
5.09. Have there been any infestations (e.g. by insects/remoold) during the storage period in this structure?	s05q09	Pest infestation (1=yes, 0=no)
5.10. Have there been any attacks by rodents during the storage period in this structure?	s05q10	Rodent attack (1=yes, 0=no)
5.12. How would you classify the losses incurred during the storage period in this structure?	s05q12	Classify losses (1=Very serious, 2=serious, 3=Negligible)
5.13. What type of pest control did you use for storing in this structure?	s05q13	Type of pest control (0=No control, 1=Sun drying, 2=Removal of infested grain, 3=Mixture with ashes and other plant materials, 4 =Smoke, 5=Pesticide/Insecticide, 97=Others)
5.16. How many weeks has the culture been stored in this structure?	s05q16	Weeks stored

The variables collected on grain storage losses, that were obtained from the laboratory analysis of the sampled grains during each of the visit, are summarized in Table 13.

Table 13: Laboratory variables assessed on the grain samples

Laboratory results	Variable	Content
First, second & third visit		
Sample Weight Grains (g)	PoidsEchantillonGrainsgrams	Total weight in g
Weight of foreign matter (g)	Poidsmatièresétrangères	Foreign materials in g
Number of undamaged grain	Nombredegrainsnonendommagés	Number of undamaged grains
Weight of undamaged grain	Poidsdegrainsnonendommagés	Weight of undamaged grains in g
Number of damaged grain	Nombregrainsendommagés	Number of damaged grains
Weight of damaged grain	Poidsdesgrainsendommagés	Weight of damaged grains in g
Percentages of foreign material	Pourcentagesmatièresétrangères	Percent of foreign materials

Variables from the EAC survey 2022/2023

Data from the first visit of the EAC survey 2022/2023 was extracted to use key indicators on the area cultivated and harvested area. Moreover, quantities of crop harvested, crop sold, crop stored, crop lost (during threshing, winnowing, drying, bagging, and transport activities), crop used as seed, consumed, given away, and used as animal feed, as well as information on pesticides, fungicide, herbicide, biopesticide and other pest controls were taken from the EAC survey 2022/2023 to be included as possible explanatory variables in the modelling exercise. Other data imported from the EAC survey 2022/2023 are institutional support received by the farm household, agricultural credits obtained by the household members, and ownership/usage of agricultural mechanical equipment by the households.

For socio-economic analysis purpose, the aggregate agricultural area at the disposal of the household, jointly with crop production and crop commercialisation were included. Outlier identification was implemented for total production and total amount earned by crop sales, considering values higher than three times the standard deviation from the average as extremes, which were replaced by the distribution average value.

Weather variables

As by the results from the literature review, it was considered highly relevant to integrate a series of atmospheric variables to evaluate the impact of weather indicators on grain storage losses. Based on the GPS information collected in the in-depth storage loss survey, data was downloaded from the Climate Data Store of the Copernicus Climate Change Service (C3S), one of the publicly available information services provided by the Copernicus Earth Observation Programme of the European Union[8].

Atmospheric data was downloaded from the ERA5-Land reanalysis dataset Following the definition of Hersbach *et al*., 2020, "[…] By optimally combining observations and models, reanalyses indeed provide consistent 'maps without gaps' of Essential Climate Variables and strive to ensure integrity and coherence in the representation of the main Earth system cycles (e.g., water, energy)". The level of detail of the data

[8] For further information, please visit https://climate.copernicus.eu/about-us and https://cds.climate.copernicus.eu

used is 0.1° x 0.1° (horizontal resolution, i.e. 9 km x 9 km resolution). In addition, the dataset included hourly information for all variables downloaded.

For weather variables, ERA5 hourly records of the following variables were selected according to geographical positions, hours and dates of each visit of Mali survey (2023):

- 2m temperature was used (temperature of air in kelvin (K) units, at 2m above the surface of land)
- 2m dewpoint temperature (temperature in kelvin (K) units, to which the air at 2m above the surface as measure of air humidity)
- total precipitation (accumulated water (liquid and/or frozen) that falls to the surface in 1 hour measured in depth meters)

Weather measurements were extracted and integrated for a 30-day period prior to each visit, and basic statistics for these three weather parameters calculated (mean, maximum, and sum for total precipitation), for geographic square blocks of 0.25° Latitude x 0.25° Longitude. Selected weather indicators are presented in Table 14.

Table 14: Selected weather variables

ERA5-Land reanalysis dataset	
mean 2m dewpoint temperature converted to degrees Celsius	mean_d2m
mean 2m temperature converted to degrees Celsius	mean_t2m
max 2m temperature converted to degrees Celsius	max_t2m
max total precipitation in depth meters	max_tp
sum 30 days total precipitation in depth meters	sum_tp

Data integration

To create a unique dataset for the analysis and modelling, databases were merged using household and crop as key links. For the households sampled, the same identification variable established by CPS for the EAC survey 2022/2023 was used in the in-depth storage loss survey. Similarly, crop IDs used in the survey were the same as the ones used in the EAC for maize and millet For weather variables, the recorded latitude and longitude and for each farm visited was linked to the corresponding geographic square blocks of 0.25° Lat x 0.25° Lon from the weather indicators in the 30-day period prior to each visit.

4.2 Storage loss modelling procedures

4.2.1 Modelling approach

Based on the wide set of data integrated on the households, farm activity, grain storage and its losses, a baseline model is to be established and build on identified explanatory variables. The models to be used require accounting for a proper specification of the distributional behaviour of the percentage of grains lost in storage. The models also need to account for the trend of the level of losses resulting from the storage losses obtained from the three visits.

To model the statistical relationship of several predictors on the level of storage grain losses, the first alternative is the use of multilevel models, also known as mixed or hierarchical models. These allow to account for the correlational structure of repeated measures. Since the recorded percentage losses are positively skewed, the natural log transformation is often used to model nonnegative, skewed dependent variables. The problem is that the reverse transformation is biased in terms of the prediction of the percentage of food lost (Cameron & Trivedi, 2010).

The use of Generalized Estimating Equations (GEE), first introduced by Liang and Zeger (Liang et al., 1986; Zegler et al., 1986; Zegler et al., 1988), also called population-averaged model (subject-specific or conditional method), is often used to analyse longitudinal and other correlated response data, particularly if responses are binary or counts. This methodology is an extension of the generalized linear model using the quasi-likelihood approach (Hanley et al., 2003; Twisk, 2003).

The first GEE tested for this research was the log normal model, using the natural logarithm as the link function and the normal distribution family, including alternatively the exchangeable and the 1st order autoregressive correlation structures. The quasi-likelihood approach solves the possible bias for the predictions, but the base models showed convergence difficulties for the parameter estimation. The second GEE tested was the Poisson model, using the natural logarithm as the link function and the Poisson distribution family GEE model. The response variable is strictly positive and has a right-skewed distribution; it looks as a "count-like" continuous variable with values commonly under 10%, and there is no reason to consider that the quasi-likelihood estimation procedure depends on the assumption of mean-variance equality. All these considerations enable the use of the Poisson distribution family. The analysis of the correlation structures showed that both structures can be used with no significant differences between models, so the exchangeable correlation structure was selected for simplicity. Therefore, the final selected family of models used was Poisson GEE with exchangeable correlation structure. Loss contributions of selected variables are estimated using coefficients β_i expressed as Incidence Rate Ratios ($\mathrm{IRR} = \mathrm{e}^{\beta_\mathrm{i}}$), interpreted as a multiplicative contribution to the percent loss, and can be interpreted also as a percentage of contribution to the percent loss. In some cases, grain storage loss estimates are presented in trend plots.

4.2.2 Steps to build the model

Initially, a descriptive analysis was developed for declared information on harvest and grain storage quantities. A screening procedure was used to identify main storage loss determinants, based on exploratory graphical correlation analysis, and regression trees (CART).

Based on these, the second step focused on testing several exploratory models considering loss drivers by groups, arranged by socio-economic and regional factors, reported storage events in each visit, storage facility characteristics and basic climate variables. This analysis was extended to include EAC survey 2022/2023 predictors related to advisory support and problems, the access to credits, the use of equipment owned or rented, the use of insecticide or herbicide and the declared loss percentages in post-harvest operations prior to storage.

Finally, a manual model selection procedure was implemented with support of subject matter experts to finetune the theoretical driven variable selection from all exploratory models. This step concluded in the finalized baseline models for each crop, first based on the in-depth storage loss survey and its potential predictors and weather variables, and second, incorporating EAC survey 2022/2023 selected predictors for obtaining the base integrated models. These base integrated models were then reduced to a minimal number of predictors to avoid multicollinearity and overfitting for the selected final models.

The specification test was used to evaluate if the percentage of food loss stored can be correctly predicted with the Poisson model, as a function of the selected explanatory variables.

4.3 Storage loss models

Base models consider a saturated approach including several drivers as predictor variables. Nevertheless, small sample sizes can lead to biased regression coefficients in GEE models, which can impact the reliability, validity, and generalizability of the study findings. Final models are intended to reduce information requirements, aiming at establishing a minimal number of predictors for a reliable prediction of stored loss. This procedure considered predictor variables from all sources: in-depth storage loss survey of this research, the EAC survey 2022/2023, and weather indicators. The base models using in-depth storage loss survey and its predictors and integrated weather variables are presented in appendix 3, and the base integrated models with the inclusion of the variables from the EAC survey 2022/2023 are presented in appendix 4. Final reduced models are described below, including some complementary findings from the base models.

4.3.1 Storage loss model for millet

The final reduced GEE model to estimate the percent of millet storage losses uses the logarithm link, the Poisson distribution, and an exchangeable correlation structure. The number of households included in the modelling are 145, with a total of 418 measurements for the three visits. The model's chi square test (hypothesis tested being null model = proposed model) is significant ($p<0.001$).

For establishing the relational structure of the divers to the estimated percent of millet storage losses, Table 15 presents the estimated incidence rate ratios (IRRs). In terms of variables obtained from EAC survey 2022/2023, one can see that the estimated millet storage losses is 45% less (1-0.55 in percent) in households with advisory support received through at least one institution (S7Q10). Households with their own machinery showed a 64.5% reduction in millet losses, and 83.2% reduction is estimated if they rent machinery. An increase in the number of temporary workers, by unit is related to a 11.1% reduction.

The selected drivers on storage characteristics include variables on storage facilities with walls of mud, which relates to an increase of 63.4% in the percentage of storage losses. Floor drying shows almost 2.6 times higher storage losses, more than twice of the estimated percent loss. Finally, the weather parameter selected is the mean temperature, related to an increment of 13% in millet storage losses by 1 degree Celsius of increment.

Table 15: Model estimated incidence rate ratios (IRR=e^(β_i)) for millet losses

pfl_f	IRR	Std. Err.	z	P>|z|	[95% Conf.	Interval]
S7Q10_APPUI_CONSEIL_TIC	.5502507	.1115437	-2.95	0.003	.3698366	.8186746
1.dmechown	.3546912	.1048641	-3.51	0.000	.198698	.6331509
1.dmechrent	.1679859	.0628279	-4.77	0.000	.0807082	.3496455
temp_wrkrs	.889855	.0437253	-2.37	0.018	.8081523	.9798176
age	1.000525	.0074481	0.07	0.944	.9860327	1.01523
walls						
Mud	1.634349	.3627116	2.21	0.027	1.057882	2.524947
1.floordry	2.585958	1.327966	1.85	0.064	.945157	7.075202
mean_t2m	1.131857	.0371406	3.77	0.000	1.061354	1.207043
_cons	.0963661	.1220936	-1.85	0.065	.008044	1.154455

The base model in appendix 3 presents small overestimation of regression coefficients for all predictors included in the final model. In this model, additional predictors showed other interesting potential contributions to millet storage losses.

When the structure is 1 meter high or more, the loss increases by 79%. If the respondent had experienced insecurity (S7Q18M6) there is a 76% reduction, and if encountered difficulties with weeds during harvest (S7Q18M1) there is an increase of 96% in millet loss. The mean temperature interacts with the maximum precipitation (depth metres of liquid water that falls to the surface) showing a positive relationship on millet losses as shown in Figure 3.

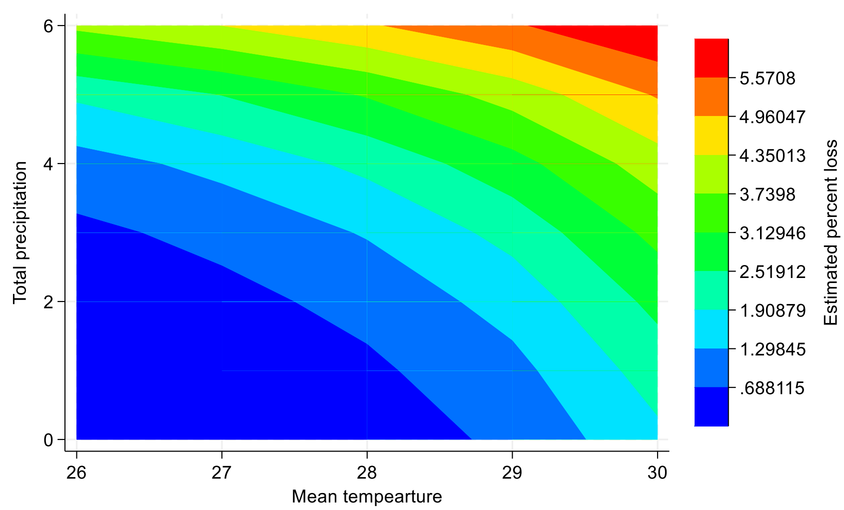

Figure 3: Weather factors interaction for millet losses

To validate the millet loss prediction capability of the final model, the specification test shows a good linear relationship between the linear predictor and the observed storage losses. Thereby, it represents a reliable way to estimate millet losses without lack of fit, as summarized in Table 16.

Table 16: Model specification test (linktest) for millet

GEE population-averaged model		Number of obs =	418
Group variable:	id2	Number of groups =	145
Link:	log	Obs per group:	
Family:	Poisson	min =	1
Correlation:	exchangeable	avg =	2.9
		max =	3
		Wald chi2(2) =	59.02
Scale parameter:	1	Prob > chi2 =	0.0000

pfl_f	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
hat	.977391	.1888084	5.18	0.000	.6073333	1.347449
hatsq	-.0232253	.1392129	-0.17	0.868	-.2960776	.249627
_cons	.0047316	.1192633	0.04	0.968	-.2290202	.2384835

This test shows that the predicted millet losses (hat) is linearly related to the observed millet losses with the model specified based on the included predictors, confirmed by the coefficient of 'hat' not being significantly different to one (linear test, $\beta hat = 1,\ p = 0.905$). The test of lack of fit also assesses weather there is a nonlinear trend of the predicted millet losses, expressed as 'hatsq'. The result shows that 'hatsq' is not significantly different from zero ($\beta hatsq = 0, p = 0.868$). This means that the selected model is a good predictor for the level millet storage losses.

4.3.2 Storage loss model for maize

Similarly, the final reduced GEE model for maize storage loss has been established. With regards to the EAC survey 2022/2023 variables, these show that households with rented machinery had a 35.9% reduction in maize storage losses, and that access to credit is related to 46.9% less maize storage loss (Table 17).

For characteristics of the storage activity, farmers' declaration of insect infestation is related to 48.7% higher losses, and rodent infestation with 61% higher losses. The use of raised racks is related to a reduction of 55.8%. Finally, the weather parameter kept in the model is the variable of maximum precipitation, with a significant increment of 6.5% in maize losses per depth metre of liquid water.

The integrated base model for maize (further detailed in appendix 3), based on these same selected predictors shows IRR coefficients with smaller standard errors. In this model, there are reductions in maize losses on the second (52.1%) and third (67.1%) visit compared to the first visit. The containers used for storage showed different levels of maize losses, "Polypropylene bags" (PP) 61.2% less losses, "On cobs" 68.4 and "Loose on the ground" with 52.7% less compared to "Jute bags".

Table 17: Model estimated incidence rate ratios (IRR=e^(β_i)) for stored maize loss

pfl_f	IRR	Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.dmechrent	.6409597	.1215982	-2.34	0.019	.4419234	.9296392
1.credit_EAC	.5310461	.0784189	-4.29	0.000	.3975908	.7092969
age	1.002844	.0046449	0.61	0.540	.9937809	1.011989
inf_insects						
Oui	1.487043	.1158597	5.09	0.000	1.276451	1.732379
inf_rodents						
Oui	1.610103	.1749348	4.38	0.000	1.301285	1.992211
1.raised_r	.4415751	.1767263	-2.04	0.041	.2015294	.967544
max_tp2	1.06486	.0279049	2.40	0.016	1.011548	1.120981
_cons	1.691814	.5358428	1.66	0.097	.9093989	3.147393

The last part of the model includes the weather variables, where the mean temperature shows a positive contribution to maize storage losses, but apparently interact negatively with the maximum precipitation on lower mean temperature (Figure 4).

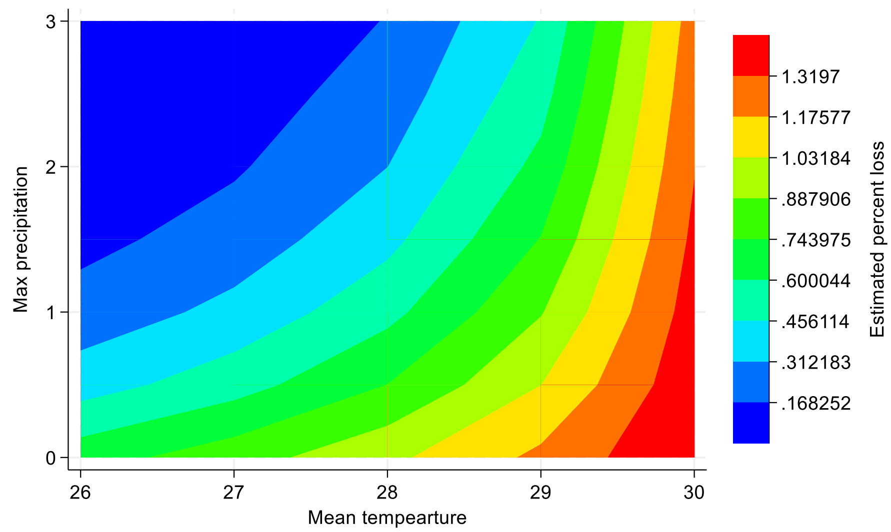

Figure 4: Weather factors interaction for maize losses

The specification test for this model shows a sufficient good linear relationship between the predicted and the observed maize storage losses, and therefore a reliable way to estimate maize losses without lack of fit, as shown in Table 18.

Table 18: Model Specification test (linktest) for maize

GEE population-averaged model		Number of obs =	434
Group variable:	id2	Number of groups =	157
Link:	log	Obs per group:	
Family:	negative binomial(k=1)	min =	1
Correlation:	exchangeable	avg =	2.8
		max =	3
		Wald chi2(2) =	35.17
Scale parameter:	1	Prob > chi2 =	0.0000

pfl_f	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
hat	.8624581	.2118329	4.07	0.000	.4472733	1.277643
hatsq	.1997278	.2163796	0.92	0.356	-.2243683	.623824
_cons	.0241646	.1008651	0.24	0.811	-.1735273	.2218566

The predicted maize storage losses (hat) have a linear correspondence; seen as being not significantly different to one (linear test of $\beta hat = 1$, $p = 0.516$). The lack of fit represented by (hatsq) is not significantly different from zero ($\beta hatsq = 0$, $p = 0.356$). This means that the selected model is a good maize loss predictor.

4.4 Lessons learnt on the modelling procedure

The modelling exercise concluded in the possibility to identify base models with sufficiently good model specifications and fit for purpose, indicating that the level of grain storage is to some extent explained by the collected and used explanatory variables. As assumed from the conceptual framework and literature review, storage-specific indicators as well as weather-related indicators were confirmed in the set of explanatory variables. Nevertheless, some variables considered key in explaining the level of losses, such as packaging during storage (whether loose on the ground, using jut bags, or hermetic bags), showed mixed results in the modelling.

Other variables were in line with the literature, for example having the presence of temporary workers was significant in reducing storage losses. Farmers' declarations on observed pests' and rodents' infestations led to higher storage losses in maize as pests are major causal factors for losses. The type of material used for the floor and the walls of the storage structure is significant for millet; for maize, the use of raised racks in storage significantly reduced losses of grains compared to grains stored on the ground. Precipitation was significant for losses in maize, higher precipitation led to higher losses as grains exposed to moisture tend to rot or develop moulds faster than grains stored in dry areas.

Some of the explanatory variables from the EAC survey were also confirmed to be significant, which can help to reduce the number of additional variables to be included in the national farm survey for storage loss estimations. Use of farm machinery both owned or rented was significant in reducing storage losses in millet and maize, this can be linked to better efficiency during harvesting and post-harvest operations. Access to extension services was also found to be significant in millet, with households that received advice on post-harvest management having less losses as compared to households that had no advice. In maize storage, access to credit was significant in reducing losses, this could be a proxy for other variables as having more available income in the household can help the household invest in better storage infrastructure and other technologies that reduce losses.

A relevant limitation has been the sample size of the survey. While there were several variables showing relevance for explaining the level of storage grain losses (as by the models presented in annex 3), the number of explanatory variables had to be reduced in the final model. With a total number of 157 households for maize and 147 household with millet used in the modelling exercise, only a handful of explanatory variables could be considered due to the small sample size and resulting problems of multicollinearity.

On the modelling approach, it seems that the use of Generalized Estimation Equation (GEE) based on Poisson models using the natural logarithm shows sufficiently good model specifications. When losses are measured based on laboratory analyses, the use of Poisson model is a good alternative to estimate storage losses given that it is based on proportions or percentages, which in the case of grain storage losses are relatively small proportions. If these small percentages are dispersed over a small range of variation, the estimated percent might be proportional to the mean proportion. Nevertheless, if extreme values are present, these may imply overdispersion. In this regard, the quantity (grain weight) of grain losses, can be expressed as an incidence rate in a Poisson model, where the total stored amount can be considered as exposure term, which in turn could result in a procedure to estimate grain loss considering adjustments based on the real stored quantity.

Another lesson learnt is on the effort to combine diverse data sources to complement the set of explanatory variables, which successfully added relevant explanatory variables. The efforts made to use variables already collected EAC survey 2022/2023 shows that relevant explanatory variables are already produced in the country. Other data is publicly available, as it was the case of weather sources like ERA5, representing a cost-efficient way to obtain data on key drivers of storage losses.

5. Conclusions and way forward

Overall, the research generated several key insights and valuable lessons for both methodological and practical applications. Notably, it identified a robust set of explanatory variables that can effectively predict grain losses during on-farm storage and validated the data collection methods and instruments used in the study. The integration of variables from the 2022/2023 EAC survey, combined with weather data from auxiliary sources—made possible by the availability of georeferenced farm locations—further enhanced the accuracy, efficiency, and optimization of the data collection process.

These findings underscore the potential for developing indirect estimation models of storage losses that can be applied more broadly. Specifically, recommendations can be derived for incorporating key explanatory variables into national agricultural surveys, allowing for the estimation of storage losses even in survey rounds where detailed modules on storage losses are not included. For instance, evidence from the pilot survey in Mali highlighted the significant impact of storage structure type (i.e. type of walls, raised rackets, packaging material) and storage conditions and drying methods (i.e. floor drying) on the extent of grain losses. These variables, therefore, present themselves as strong candidates for inclusion in future iterations of the EAC survey (as well as other farm surveys), supporting loss modelling through indirect methods.

The modeling exercise revealed a wide array of explanatory variables showing statistical significance, indicating the multifaceted nature of storage losses. However, only a limited number of these variables could be included in the baseline model due to constraints related to sample size. Nevertheless, the developed models still achieved a satisfactory level of predictive accuracy for both crops examined. That said, the predictive power and robustness of the models could be further improved by increasing the sample size, which would allow for the inclusion of a greater number of explanatory variables. As a general rule of thumb in such modeling efforts, roughly 50 observations are required per explanatory variable to ensure reliable and unbiased estimates. This consideration should guide future survey design and sampling strategies when planning to build or extend such predictive models.

Beyond the modeling component, the in-depth data collection on storage losses provided a rich set of indicators on storage characteristics, practices, and dynamics. These indicators hold high relevance for policymakers, particularly in countries where on-farm storage is a recognized priority area for agricultural development and post-harvest loss reduction. The use of physical measurements to estimate storage losses—especially losses due to pest infestation—significantly improved the precision of the loss estimates. However, this approach proved to be both resource-intensive and operationally challenging in field conditions. Therefore, for future implementation, it is recommended that capacity be built within National Statistical Offices or Ministries of Agriculture to conduct grain sample analysis directly in the field. Alternatively, lighter methods that yield comparable precision—such as visual assessment scales—could be considered as cost-effective substitutes.

In addition, it is advisable to conduct qualitative studies on storage practices and household-level behaviors prior to deploying an in-depth storage loss survey module. Such qualitative preparatory work would help refine the design and wording of questions related to storage characteristics, storage activities, and storage duration, thereby improving the overall quality and relevance of the collected data. Furthermore, lessons from the Mali pilot suggest that the timing of survey visits should be better aligned with the crop calendar. Ideally, the first visit should take place closer to the harvest period, and the final visit should coincide with the end of the storage season to capture the full arc of storage dynamics and loss progression.

Despite these limitations, the storage loss estimates produced in this study exhibited relatively low standard errors, indicating that the data and methodology were of sufficient quality to support meaningful analysis. The primary research objective—to identify and specify baseline models using data from a single survey round—was successfully achieved. This included the selection and validation of key explanatory variables for grain storage losses. Looking ahead, a second survey round would be necessary to complete and further refine the storage loss modelling process. Such a follow-up would allow researchers to capture interannual variations in storage losses and rigorously test the feasibility of indirect estimation approaches based on variables embedded in standard agricultural surveys.

By addressing both methodological and operational aspects of measuring and modelling on-farm storage losses, this research lays the groundwork for more scalable and sustainable approaches to post-harvest data collection and policy design.

References

50x2030 Initiative. 2021. *A guide to the 50x2030 data collection approach: Questionnaire design*. 50x0203 Technical paper series #2. https://www.50x2030.org/sites/default/files/resources/documents/2022-08/50x2030%20TECHNICAL%20PAPER%20SERIES%20%232_July22_EN.pdf

Abdoulaye, T., Ainembabazi, J., Alexander, C., Baributsa, D., Kadjo, D., Moussa, B., Omotilewa, T., Ricker-Gilbert, J., Feleke, S.2016. *Postharvest Loss of Maize and Grain Legumes in Sub-Saharan Africa: Insights from Household Survey Data in Seven Countries*. https://doi.org/10.13140/RG.2.1.2327.8960.

Abass, A.B., Ndunguru, G., Mamiro, P., Alenkhec, B., Mlingia, N., Bekunda, M. 2013. *Post-harvest food losses in a maize-based farming system of semi-arid savannah area of Tanzania*. Journal of Stored Products Research, Volume 57, Pages 49-57. ISSN 0022-474X. https://doi.org/10.1016/j.jspr.2013.12.004.

Adejumo, B.A., Raji, A.O., 2007. Technical appraisal of grain storage systems in the Nigerian Sudan savannah. Agricultural Engineering International, International Commission of Agricultural Engineering (CIGR) Ejournal, No. 11, Vol. IX, September 2007.https://hdl.handle.net/1813/10669

Baidhe, E.; Clementson, C.L.; Senyah, J.; Hammed, A. 2024. Appraisal of Post-Harvest Drying and Storage Operations in Africa: Perspectives on Enhancing Grain Quality. AgriEngineering 2024, 6, 3030–3057. https://doi.org/10.3390/ agriengineering6030174

Baributsa, D., Njoroge, A. 2020. The use and profitability of hermetic technologies for grain storage among smallholder farmers in eastern Kenya. Journal of Stored Products Research. https://doi.org/87. 101618. 10.1016/j.jspr.2020.101618.

Boxall, R.A. 1986. A Critical Review of the Methodology for Assessing Farm-Level Grain Losses after Harvest; Nature Resource Institute, University of Greenwich: London, UK, 1986; ISBN 0 85954-208-4. https://gala.gre.ac.uk/id/eprint/10793/

Cameron, A.C., and Trivedi, P.K. 2010. *Microeconometrics Using Stata*. Rev. ed. College Station, TX: Stata Press, p. 103. Revised Edition, StataCorp LLC, https://EconPapers.repec.org/RePEc:tsj:spbook:musr.

Compton, J., Floyd, S., Ofosu, A. and Agbo, B. 1998. The Modified Count and Weigh Method: An Improved Procedure for Assessing Weight Loss in Stored Maize Cobs. *Journal of Stored Products Research, 34 (4). pp. 277-285. ISSN 0022-474X.* https://doi.org/10.1016/S0022-474X(98)00009-5

Compton, J.A.F., Sherington, J. 1999. Rapid assessment methods for stored maize cobs: weight losses due to insect pests, *Journal of Stored Products Research, Volume 35, Issue 1, Pages 77-87*, ISSN 0022-474X, https://doi.org/10.1016/S0022-474X(98)00012-5.

De Groote, H., Kimenju, S., Likhayo, P., Kanampiu, F., Tefera, T., Hellin, J. 2013. Effectiveness of hermetic systems in controlling maize storage pests in Kenya. *Journal of Stored Products Research. 53. 27–36.* https://doi.org/10.1016/j.jspr.2013.01.001.

De Groote, H., Muteti, F.N., Bruce, A.Y. 2023. On-farm storage loss estimates of maize in Kenya using community survey methods, *Journal of Stored Products Research, Volume 102,* https://doi.org/10.1016/j.jspr.2023.102107.

FAO. 2017. *Field Test Report on the Estimation of Crop Yields and Post-Harvest Losses in Ghana*. Rome. https://openknowledge.fao.org/handle/20.500.14283/ca6429en

FAO and Ethiopian Statistics Service (ESS). 2023. *Report on pre- and post-harvest crop losses pilot survey (2021–2022).* Rome and Addis Ababa. https://doi.org/10.4060/cc3937en

Freitas, R.S., Faroni, L.R.A., Sousa, A.H. 2016. Hermetic storage for control of common bean weevil, Acanthoscelides obtectus (Say), Journal of Stored Products Research, Volume 66, https://doi.org/10.1016/j.jspr.2015.12.004.

GSARS (Global Strategy to improve Agricultural and Rural Statistics). 2018. *Guidelines on the measurement of harvest and post-harvest losses*. Rome. https://openknowledge.fao.org/handle/20.500.14283/ca6396en

Hanley, J.A., Negassa, A., Edwardes, M.D., Forrester, J.E.2003. Statistical analysis of correlated data using generalized estimating equations: an orientation, *American Journal of Epidemiology 157(4), 364–375.* https://doi.org/10.1093/aje/kwf215.

Harris, K.L. and Lindblad, C.J.1978. Post-Harvest Grain Loss Assessment Methods. *American Association of Cereal Chemist, Minneapolis, pp. 193.*

Hersbach, H., Bell, B., Berrisford, P. 2020. The ERA5 global reanalysis. QJR Meteorol Soc. 2020;146:1999–2049. https://doi.org/10.1002/qj.3803

Hodges, R.J., Buzby, J.C., Bennett, B., 2011. Postharvest losses and waste in developed and less developed countries: opportunities to improve resource use. *The Journal of Agricultural Science*. 2011;149(S1):37-45. https://doi.org/10.1017/S0021859610000936

Hodges, R.J., Bernard, M., Knipschild, H., Rembold, F. 2014. *APHLIS - Postharvest cereal losses in Sub-Saharan Africa, their estimation, assessment and reduction*. https://doi.org/10.2788/19582.

Hodges, R.J., Buzby, J., Bennett, B.2011. Postharvest losses and waste in developed and less developed countries: Opportunities to improve resource use. *The Journal of Agricultural Science. 149. 37 - 45.* https://doi.org/10.1017/S0021859610000936.

Hsiao, C. 2014. Analysis of Panel Data. 3rd ed. Cambridge University Press; https://doi.org/10.1017/CBO9781139839327

Kadjo, D., Ricker-Gilbert, J., Abdoulaye, T., Shively, G. and Baco, M.N. 2018. Storage losses, liquidity constraints, and maize storage decisions in Benin. *Agricultural Economics, 49: 435-454.* https://doi.org/10.1111/agec.12427

Kebe, M. 2016. *Developing cost-effective statistical methods for measuring post-harvest losses in developing countries*. ICAS VII Seventh International Conference on Agricultural Statistics I Rome 24-26 October 2016. https://doi.org/10.1481/icasVII.2016.b15d

Kiaya, V. 2014. Post-Harvest Losses and Strategies to Reduce them. Action Contre la Faim (ACF). https://www.actioncontrelafaim.org/wp-content/uploads/2018/01/technical_paper_phl__.pdf

Kumar, D., Kalita, P. 2017. Reducing Postharvest Losses during Storage of Grain Crops to Strengthen Food Security in Developing Countries. *Foods.* 2017 Jan 15;6(1):8. https://doi.org/10.3390/foods6010008.

Liang, K.-Y., Zeger, S.L. 1986. Longitudinal data analysis using generalized linear models. Biometrika, 73(1), 13-22. https://doi.org/10.2307/2336267

Manandhar, A., Milindi, P., Shah, A. 2018. An Overview of the Post-Harvest Grain Storage Practices of Smallholder Farmers in Developing Countries. *Agriculture*, *8*(4), 57. https://doi.org/10.3390/agriculture8040057

Memon, M., Ting, H., Cheah, J.-H., Ramayah, T., Chuah, F., Cham, T-H. 2020. Sample Size for Survey Research: Review and Recommendations. Journal of Applied Structural Equation Modeling. 4. i-xx. https://doi.org/10.47263/JASEM.4(2)01 .

Mendesil, E., Kuyu, C.G., Anderson, P. 2022.Effects of storage in triple-layer hermetic bags on stored field pea grain quality and infestation by the pea weevil, Bruchus pisorum L. (Coleoptera: Bruchidae), Journal of Stored Products Research, Volume 95. https://doi.org/10.1016/j.jspr.2021.101919.

Mlambo et al., 2017. Mlambo, Shaw & Mvumi, Brighton & Stathers, Tanya & Mubayiwa, Macdonald & Nyabako, Tinashe. (2017). Field efficacy of hermetic and other maize grain storage options under smallholder farmer management. Crop Protection. 98. https://doi.org/10.1016/j.cropro.2017.04.001.

Ndegwa, M.K., De Groote, H., Gitonga, Z.M., Bruce, A.Y., Effectiveness and economics of hermetic bags for maize storage: Results of a randomized controlled trial in Kenya, Crop Protection, Volume 90, https://doi.org/10.1016/j.cropro.2016.08.007.

Ng'ang'a, J., Mutungi, C., Imathiu, S., & Affognon, H. (2016). Low permeability triple-layer plastic bags prevent losses of maize caused by insects in rural on-farm stores. *Food Security, 8* (3), 621-633. https://doi.org/10.1007/s12571-016-0567-9

Ngwenyama, P., Mvumi, B.M., Nyanga, L.K., Stathers, T.E., Siziba, S. 2020. Comparative performance of five hermetic bag brands during on-farm smallholder cowpea (Vigna unguiculata L.Walp) storage, Journal of Stored Products Research, Volume 88, https://doi.org/10.1016/j.jspr.2020.101658

Nukenine, E.N.2010. Stored product protection in Africa: Past, present and future. 10th International Working Conference on Stored Product Protection. https://doi.org/10.5073/jka.2010.425.177

Ognakossan, K.E., Agbeko Kodjo, T. & Lamboni, L.Y., Hell, K. 2013. Post-harvest insect infestation in maize grain stored in woven polypropylene and in hermetic bags. International Journal of Tropical Insect Science. 33. 71–81. https://doi.org/10.1017/S1742758412000458.

Ratnadas A, Berte, S., Diarra, D., and Cisse, B.,1990. Insect losses on sorghum stored in selected Malian villages, with particular emphasis on varietal differences in grain resistance. Stored product protection. Volume 2. pp. 953-959. International Working Conference on Stored-product Protection. 6, Canberra, Australie, 17 April 1994/23 April 1994. https://agritrop.cirad.fr/387618/

Wollburg, P., Tiberti, M., Zezza, A., 2021.Recall length and measurement error in agricultural surveys, Food Policy, Volume 100, https://doi.org/10.1016/j.foodpol.2020.102003.

Twisk, J., 2003 Applied Longitudinal Data Analysis for Epidemiology: A Practical Guide. Cambridge University Press, New York, U.S.A. 301 pp. https://assets.cambridge.org/97805218/19763/frontmatter/9780521819763_frontmatter.pdf

Sattar. M, Mueen-u-Din, A. M., Ali, L., Waqar, M. Q., Ali, M.A. and Khalid, L. 2015. Grain Losses of Wheat as Affected by Different Harvesting and Threshing Techniques. International Journal of Research in Agriculture and Forestry, 2: 20-26. https://www.ijraf.org/pdf/v2-i6/4.pdf

Shee, A., Mayanja, S., Simba, E., Stathers, T., Bechoff, A., Bennett, B..2019. Determinants of postharvest losses along smallholder producers' maize and sweetpotato value chains: an ordered Probit analysis. Food Security. 11. https://doi.org/10.1007/s12571-019-00949-4.

Zeger, S.L., Liang, K.-Y. 1986. Longitudinal data analysis for discrete and continuous 0utcomes. *Biometrics, 42(1), 121-130*. https://doi.org/10.2307/2531248

Zeger, S.L., Liang, K.-Y., Albert, P.S. 1988 Models for Longitudinal Data: A Generalized Estimating Equation Approach. *Biometrics, 44(4), 1049-1060*. https://doi.org/10.2307/2531734

Additional related reading:

Afrique Verte. 2004. *Fiche Technique sur le Commerce Mil/Sorgho au Mali*. La cellule Capitalisation. https://www.afriqueverte.org/r2_public/media/fck/File/Documentation/Outils_information/commerce-mil-sorgho-mali.pdf

Aulakh, J.; Regmi, A. 2013. *Post-Harvest Food Losses Estimation-Development of Consistent Methodology*. In Proceedings of the Agricultural & Applied Economics Association's 2013 AAEA & CAES Joint Annual Meeting, Washington, DC, USA, 4–6 August 2013. https://www.fao.org/fileadmin/templates/ess/documents/meetings_and_workshops/GS_SAC_2013/Improving_methods_for_estimating_post_harvest_losses/Final_PHLs_Estimation_6-13-13.pdf

Cellule Nationale CEDEAO. 2010. *Plan National D'investissement Prioritaire Dans Le Secteur Agricole Au Mali- 2011-2015*. Rapport PNIP-SA. CAADP Country Investment Plans. https://faolex.fao.org/docs/pdf/mli144900.pdf

Chen, X., Wu, L., Shan, L., & Zang, Q. (2018). Main Factors Affecting Post-Harvest Grain Loss during the Sales Process: A Survey in Nine Provinces of China. Sustainability, 10(3), 661. https://doi.org/10.3390/su10030661

CSA. 2007. *Rapport de Synthèse des Missions de Suivi-Evaluation des Banques de Céréales Installées en 2005 par le Commissariat à la Sécurité Alimentaire*. https://reliefweb.int/attachments/9f44ed26-4731-327e-aba7-acd58eeb5721/C724503F44435C5EC125735B0031AA24-Rapport_Complet.pdf

Dembele, N.N., Diarra, S.B., 2007. Rôle de Banques de céréales dans le dispositif de stockage de céréales au Mali, Fo*od Security Collaborative Working Papers 57058, Michigan State University, Department of Agricultural, Food, and Resource Economics*. https://doi.org/10.22004/ag.econ.57058

DNA. DNPIA. DNSI. R*ecensement général de l'Agriculture (RGA)- Campagne Agricole 2004-2005. Résultats définitifs - Volume II*: *Rapport Détaillé*. FAO Project GCP/INT/903/FR. Appui au Programme de Renforcement des Systèmes d'Information et de Statistiques Rurales en Afrique. FAO https://www.fao.org/fileadmin/templates/ess/ess_test_folder/World_Census_Agriculture/Country_info_2000/Eliman_files/REP2_MLI_FRE_2005.pdf

Fall, A. 2011. *Etude Sur la Chaine de Valeur Mil/Sorgho in Mali.* Initiatives Intégrées Pour la Croissance Économique in Mali (IICEM), Agence Américaine pour le Développement International, USAID. Report. https://pdf.usaid.gov/pdf_docs/pa00hnwp.pdf

FAO. 2011. *Global Food Losses and Food Waste—Extent, Causes and Prevention*; Rome, Italy. https://openknowledge.fao.org/handle/20.500.14283/i2697e

FAO. 2013. *Cadre De Programmation Pays (CPP) 2013-2017. FAO - Mali*. Rome. https://openknowledge.fao.org/handle/20.500.14283/bl902f

FAO. 2020a. *Guidelines on the measurement of harvest and post-harvest losses. Estimation of crop harvest and post-harvest losses in Malawi. Maize, rice and groundnuts*. *Field Test Report.* Rome. https://openknowledge.fao.org/handle/20.500.14283/cb1562en

FAO. 2020b. *Guidelines on the measurement of harvest and post-harvest losses. Estimation of crop harvest and post-harvest losses in Zimbabwe. Field Test Report*. Rome. https://openknowledge.fao.org/handle/20.500.14283/cb1554en

INSTAT (Institut National de la Statistique). 2019. *Etude d'évaluation de la qualité des statistiques agricoles et propositions d'amélioration*. RAPPORT. instat-mali.org/laravel-filemanager/files/shares/eq/rapport-eva-stat-agric_eq.pdf

Kaminski, J., Christiaensen, J. 2014. *Post-Harvest Loss in Sub-Saharan Africa What Do Farmers Say?* The World Bank Africa Region Office of the Chief Economist. World Bank. https://documents1.worldbank.org/curated/en/782981468008454485/pdf/Post-harvest-loss-in-Sub-Saharan-Africa-what-do-farmers-say.pdf

Kitinoja, L., Saran, S., Roy, S., Kader, A. 2011. Postharvest technology for developing countries: Challenges and opportunities in research, outreach and advocacy. *Journal of the science of food and agriculture. 91.* https://doi.org/597-603. 10.1002/jsfa.4295.

Manful, J., Médard, M., Kromah, A., Paul Houssou, P. *Manuel de formation sur l'amélioration des technologies post récolte du riz (draft).* CORAF/USAID/SONGHAI. https://www.doc-developpement-durable.org/file/Culture/Culture-plantes-alimentaires/FICHES_PLANTES/riz/ManuelsurleRizfinal.pdf

Mesterházy, A., Oláh, J., Popp, J. 2020. Losses in the Grain Supply Chain: Causes and Solutions. *Sustainability* 12, no. 6: 2342. https://doi.org/10.3390/su12062342

Ministère du Développement Rural du Mali. 2014. *Plan National d'investissement dans le Secteur Agricole*. 2014. https://faolex.fao.org/docs/pdf/mli198371.pdf

Ramalho, E.A., Ramalho, J. Murteira, J. 2011. Alternative estimating and testing empirical strategies for fractional regression models. Journal of Economic Surveys, 25(1), 19-68. https://doi.org/10.1111/j.1467-6419.2009.00602.x

Tossou, H.S., Manda, S., Jean, K., Assefa, W. 2020. *L'avenir de l'agriculture au Mali: 2030-2063.* Volume I, Étude de cas : Défis et Opportunités pour les projets financés par le FIDA, réalisées 2020. https://www.inter-reseaux.org/wp-content/uploads/Mali_IFAD-Futur-de-lagri.pdf

USAID. Cereal Value Chain (CVC) activity in Mali. Final Performance Evaluation.2020.

Wu, L., Hu, O., Wang J., Zhu, D. 2017. Empirical analysis of the main factors influencing rice harvest losses based on sampling survey data of ten provinces in China, China Agricultural *Economic Review,* *Emerald Group Publishing Limited, vol. 9(2), pages 287-302, May*. https://doi.org/10.1108/CAER-03-2016-0036

Liu, X. Li, B., Shen, D., Cao, J. Mao, B. 2017. Analysis of Grain Storage Loss Based on Decision Tree Algorithm, Procedia Computer Science, Volume 122, Pages 130-137, ISSN 1877-0509. https://doi.org/10.1016/j.procs.2017.11.351.

ANNEX 1

Table 19: Total households and percentage per region by harvest duration and techniques used

Cultivation methods	Millet			Maize		
	Total	Ségou	Sikasso	Total	Ségou	Sikasso
Seed variety						
Traditional	97.9%	96.5%	100.0%	93.6%	94.8%	92.7%
Modern/Improved	2.1%	3.5%	0.0%	6.4%	5.2%	7.3%
Harvest duration (avg, days)	9.7	12.7	5.4	7.1	5.0	8.6
Harvest range						
<2 days	8.7%	2.1%	18.2%	12.8%	12.4%	13.1%
2 to 5 days	29.5%	13.4%	52.5%	47.4%	53.6%	43.1%
6 to 10 days	18.7%	21.1%	15.2%	19.2%	25.8%	14.6%
10 days or more	43.2%	63.4%	14.1%	20.5%	8.2%	29.2%
Households reporting an early harvest	18.7%	28.2%	5.1%	32.9%	37.1%	29.9%
Households reporting a late harvest	8.7%	14.1%	1.0%	5.1%	6.2%	4.4%
Households reporting rain during the harvest	22.8%	16.2%	32.3%	22.6%	27.8%	19.0%
Harvest method						
Manual/Tradition	98.3%	98.6%	98.0%	94.9%	92.8%	96.4%
Machinary/Modern	1.7%	1.4%	2.0%	5.1%	7.2%	3.6%

Table 20: Total households and percentage by region according to storage structure characteristics

Storage structure characteristics		Région	
	Total	Ségou	Sikasso
Total number of storage facilities	428	201	227
Average number of storage facilities	1.5	1.4	1.6
Roof/cover			
No roof/cover	4.0%	2.6%	5.3%
Grass cover	47.5%	41.8%	52.4%
Palm leaf cover	1.4%	1.0%	1.8%
Plastic cover	0.2%	0.5%	0.0%
Metal	35.0%	39.3%	31.3%
Other	11.8%	14.8%	9.3%
Walls			
Banco bricks	61.7%	67.9%	56.4%
Woven basket	0.9%	2.0%	0.0%
Mud	30.0%	15.8%	42.3%
Cot	0.0%	0.0%	0.0%
Open wall	0.0%	0.0%	0.0%
Other	7.3%	14.3%	1.3%
Floor			

Concrete	13.5%	20.9%	7.0%
Earth	67.6%	54.1%	79.3%
Woven basket	0.2%	0.5%	0.0%
Wood	16.1%	24.0%	9.3%
Other	2.6%	0.5%	4.4%
Platform-height			
0 metres - ground level	7.1%	15.3%	0.0%
0.5 metres	8.7%	18.9%	0.0%
1 metre	0.5%	0.5%	0.4%
More than 1 metre	83.7%	65.3%	99.6%
Age of the storage facility (avg)	8.5	7.2	9.7
Age range			
< 2 years	8.3%	11.2%	5.7%
2 to 5 years	24.1%	29.6%	19.4%
6 to 10 years	31.4%	31.1%	31.7%
10 years or more	36.2%	28.1%	43.2%

Table 21: Number and percentage of households by region and cost and storage capacity

Storage facility cost		**Région**	
	Total	**Ségou**	**Sikasso**
Storage facility construction cost (avg, FCFA)	87,671.87	40,384.94	131,018.23
Amount range			
<50,000	62.5%	69.3%	56.3%
50,000 to 150,000	27.2%	29.0%	25.5%
150,000 or more	10.3%	1.7%	18.2%
Storage facility maximum capacity (avg, kg)	16,222.14	6,052.41	25,003.06
Capacity range			
< 3,000 kg	15.6%	15.3%	15.9%
3,000 to 5,000 kg	29.1%	36.2%	22.9%
5,000 to 10,000 kg	33.1%	34.2%	32.2%
10,000 kg or more	22.2%	14.3%	29.1%
Operating mode			
Collective / shared	20.8%	41.8%	2.6%
Individual	79.2%	58.2%	97.4%
Facility location			
Within the residency	10.6%	1.0%	18.9%
Inside the concession	87.0%	96.4%	78.9%
Other	2.4%	2.6%	2.2%
Quantity stored (per storage facility)			
Millet			
Stored quantity (avg)	2,486.10	2,901.34	1,811.33

Stored quantity range			
<500 kg	23.0%	13.5%	38.5%
500 to 1,000 kg	18.7%	14.7%	25.0%
1,000 to 2,000 kg	21.0%	23.1%	17.7%
2,000 to 4,000 kg	27.0%	35.3%	13.5%
4,000 kg or more	10.3%	13.5%	5.2%
Maize			
Stored quantity (avg)	2,244.77	793.76	2,856.26
Stored quantity range			
<500 kg	27.1%	55.9%	15.0%
500 to 1,000 kg	17.1%	20.3%	15.7%
1,000 to 2,000 kg	16.6%	13.6%	17.9%
2,000 to 4,000 kg	19.6%	6.8%	25.0%
4,000 kg or more	19.6%	3.4%	26.4%

Table 22: Number and percentage of households that suffered attacks, by region and collection phase

Types of attack and infestations	**Phase**	**Millet**			**Maize**		
		Total	**Ségou**	**Sikasso**	**Total**	**Ségou**	**Sikasso**
Households having suffered infestations	1	20.3%	27.1%	10.4%	14.5%	20.7%	11.7%
	2	36.2%	53.6%	10.5%	25.9%	46.6%	16.5%
	3	43.4%	67.2%	7.7%	27.1%	48.2%	17.6%
Households having suffered rodent attacks	1	12.3%	15.0%	8.3%	8.6%	15.5%	5.5%
	2	13.2%	19.3%	4.2%	9.2%	10.3%	8.7%
	3	15.8%	23.4%	4.4%	5.0%	10.7%	2.4%
Households having suffered losses	1	17.8%	25.7%	6.3%	10.8%	22.4%	5.5%
	2	25.5%	40.7%	3.2%	9.7%	22.4%	3.9%
	3	21.5%	35.0%	1.1%	8.3%	21.4%	2.4%
Loss magnitude							
Very serious	1	11.9%	0.0%	83.3%	10.0%	0.0%	28.6%
	2	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	3	0.0%	0.0%	0.0%	6.7%	0.0%	33.3%
Serious	1	4.8%	5.6%	0.0%	15.0%	15.4%	14.3%
	2	5.0%	3.5%	33.3%	22.2%	15.4%	40.0%
	3	8.2%	8.3%	0.0%	6.7%	0.0%	33.3%
Negligible	1	83.3%	94.4%	16.7%	75.0%	84.6%	57.1%
	2	95.0%	96.5%	66.7%	77.8%	84.6%	60.0%
	3	91.8%	91.7%	100.0%	86.7%	100.0%	33.3%

Table 23: Number and percentage of households by pest checks by region and visit

Pest control	Millet			Maize		
	Total	Ségou	Sikasso	Total	Ségou	Sikasso
Visit 1						
Pest control						
No control	30.5%	18.6%	47.9%	38.7%	20.7%	46.9%
Sun drying	10.2%	17.1%	0.0%	4.3%	13.8%	0.0%
Removal of infested grain and destruction	0.0%	0.0%	0.0%	0.5%	0.0%	0.8%
Mixed with ash and other vegetable matter	5.5%	9.3%	0.0%	4.3%	12.1%	0.8%
Smoking	0.4%	0.7%	0.0%	0.0%	0.0%	0.0%
Pesticide/Insecticide	47.0%	45.0%	50.0%	50.0%	48.3%	50.8%
Other	6.4%	9.3%	2.1%	2.2%	5.2%	0.8%
Visit 2						
Pest control						
No control	55.3%	36.4%	83.2%	74.1%	46.6%	86.6%
Sun drying	0.4%	0.7%	0.0%	0.0%	0.0%	0.0%
Removal of infested grain and destruction	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Mixed with ash and other vegetable matter	7.7%	12.9%	0.0%	3.8%	12.1%	0.0%
Smoking	0.4%	0.0%	1.1%	0.5%	0.0%	0.8%
Pesticide/Insecticide	30.2%	41.4%	13.7%	16.8%	29.3%	11.0%
Other	6.0%	8.6%	2.1%	4.9%	12.1%	1.6%
Visit 3						
Pest control						
No control	72.8%	64.2%	85.7%	83.4%	82.1%	84.0%
Sun drying	0.4%	0.7%	0.0%	1.1%	0.0%	1.6%
Removal of infested grain and destruction	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Mixed with ash and other vegetable matter	0.4%	0.7%	0.0%	0.0%	0.0%	0.0%
Smoking	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Pesticide/Insecticide	21.9%	32.1%	6.6%	10.5%	17.9%	7.2%
Other	4.4%	2.2%	7.7%	5.0%	0.0%	7.2%

ANNEX 2

Base models using Mali 2023 predictors

Base models for percentage loss of stored millet and maize, using Mali 2023 predictors and weather variables are presented. Initial Stata output of all models is omitted, only estimated coefficients expressed as rate ratios (RR), interpreted as a multiplicative contribution to the percent loss or as a percentage of contribution to the percent loss are presented.

Base model for millet

For the establishment of the relational structure of food loss drivers to the estimated percent of millet loss, table 1 presents the estimated rate ratios.

Table 24: Model estimated rate ratios (IRR=e^β) for millet losses

pfl_f	IRR	Std. Err.	z	P>|z|	[95% Conf.	Interval]
visit						
2	.2568635	.1316162	-2.65	0.008	.0940912	.7012227
3	.0247928	.0233199	-3.93	0.000	.0039237	.1566603
FMPc	1.029752	.0065921	4.58	0.000	1.016913	1.042754
harv_early	.6789201	.1390305	-1.89	0.059	.4544729	1.014213
dryingtime	.9822801	.0102883	-1.71	0.088	.962321	1.002653
1.floordry	2.19335	.6299133	2.73	0.006	1.24925	3.850938
treshing						
Machine moderne	1.313923	.2410383	1.49	0.137	.9171019	1.882445
prac_taining	.7388107	.1762597	-1.27	0.204	.4628712	1.179251
temp_wrkrs	.9305375	.0299008	-2.24	0.025	.8737403	.9910267
container						
Loose on the ground	.8185617	.1023934	-1.60	0.109	.6405826	1.04599
walls						
Mud	1.416077	.2098586	2.35	0.019	1.059111	1.893357
credit						
Oui	1.269672	.1661053	1.83	0.068	.9825	1.640779
age	1.002697	.0043043	0.63	0.530	.9942959	1.011169
stockee_ton	1.043479	.0327804	1.35	0.175	.9811683	1.109746
mean_t2m	1.597847	.2451405	3.05	0.002	1.18289	2.15837
sum_tp2	.2346952	.3107832	-1.09	0.274	.0175123	3.145318
c.mean_t2m#c.sum_tp2	1.058216	.0455584	1.31	0.189	.9725865	1.151384
c.sum_tp2#c.sum_tp2	.9906542	.0052925	-1.76	0.079	.9803352	1.001082
cer						
BAROUELI	1.711966	.4163658	2.21	0.027	1.062857	2.757497
KOUTIALA	2.116079	.5407811	2.93	0.003	1.282332	3.491912
SIKASSO	2.73007	.8090751	3.39	0.001	1.527265	4.88015
_cons	3.51e-07	1.43e-06	-3.65	0.000	1.21e-10	.0010173

The estimated on-farm millet loss showed a reduction of 74.3% (1-0.257 in percent) of millet loss in the second visit and a reduction of 97.5% (1-0.025 in percent) in the third, in contrast to the first visit.

Foreign Materials Percentage (FMPc) observed from laboratory analyzed samples, shows that it represents 3% of millet losses ($p<0.001$).

Harvesting and manipulating procedures related to millet losses are: harvesting early with a reduction of 32.1% ($p=0.059$), drying time with an increase of 1.8% per day ($p=0.088$). Floor drying with 2.2 times higher loses, more than twice, the estimated percent loss ($p=0.006$). Threshing using modern machines does not show a significant contribution in the loss percent ($p=0.137$). An increment per unit in the number of temporary workers reduces loss by 7% ($P=0.025$). The reported participation in practical training and the total stored grains are not statistically significant contributors to the estimated percent loss ($p=0.2$ and $p=0.175$ respectively).

The option "Loose on the ground" reported as a container shows a 18.1% reduction compared to other containers ($p=0.1$). The use of mud for the walls of the storage facility contributes to an increase of 41.6% in losses ($p=0.019$). Access to credit, reported in the questionnaire, tends to increase millet losses by 27%, and the age of the household head does not significantly contribute to food loss ($p=0.53$).

The model estimated percent of millet loss in each visit showed a decreasing trend, with an 8.15% in the first visit, 2.1% in the second, and with 0.2% in the last visit, as shown in Figure 1. The model estimated percent of millet loss in the first visit shows a wide estimation interval compared to the second and third visit, this is attributable to higher randomness in the first visit, implying a big standard error for this estimate.

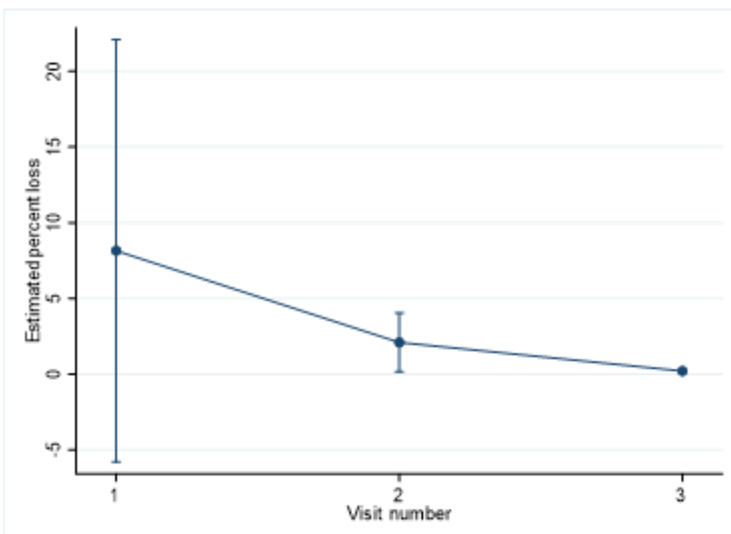

Figure 5: Estimated percent of millet loss per visit

The mean temperature is related to an increment of 1.6 times millet losses by 1 degree Celsius of increment, but its effect interacts with the previous month's total precipitation ($p=0.189$) and a curve (quadratic) effect of previous total precipitation ($p=0.079$). The contribution of these environmental factors can be understood graphically by observing the contour plot shown in figure 6.

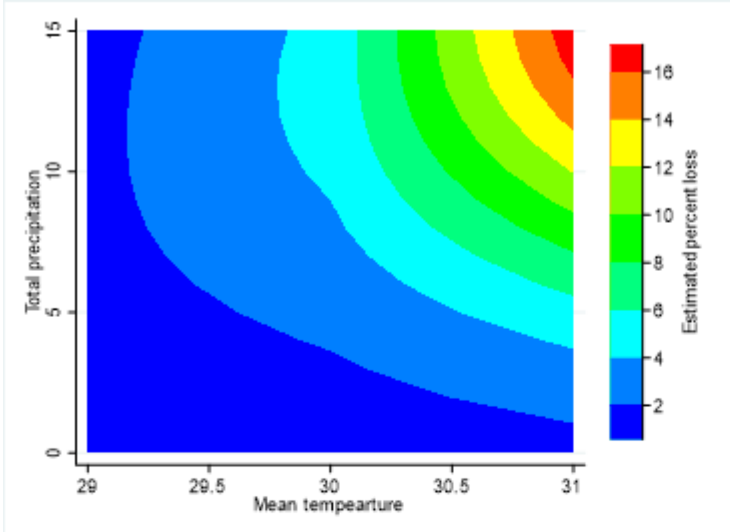

Figure 6: Weather factors interaction for millet loss

Clearly, the joint increase in temperature and precipitation represents an important cause of millet losses in storing facilities. The inclusion of the geographic division in Cercles, shows that millet losses are clearly different by region, as shown in table 25.

Table 25: Estimated Percentage of Millet Lost by Cercle

cer	Margin	Delta-method Std. Err.	[95% Conf.	Interval]
BAROUELI	.4903334	.0823177	.3289937	.6516732
BLA	.2864155	.0557983	.1770527	.3957782
KOUTIALA	.6060777	.0687153	.4713983	.7407571
SIKASSO	.7819342	.132197	.5228328	1.041035

It is also clear that BLA has the smallest estimate compared to the other three Cercles. As shown graphically in figure 7.

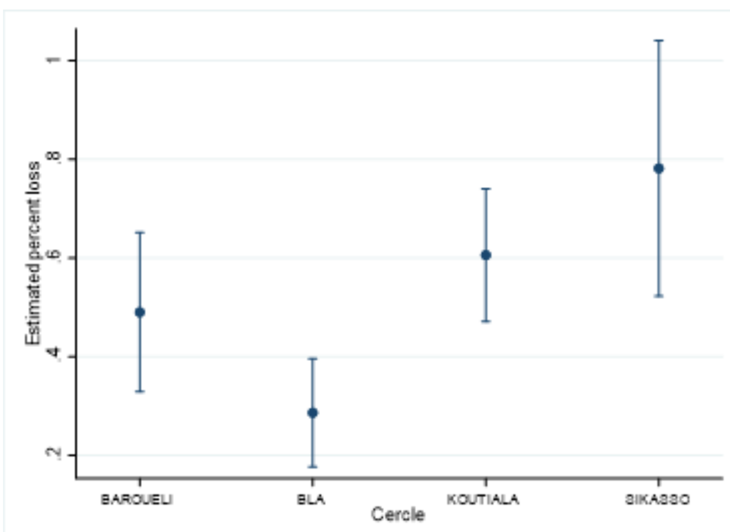

Figure 7: Estimated percentage loss of millet by district

The specification test shows a good linear relationship between the linear predictor and the observed food loss; the model is a good millet loss predictor as shown in table 26.

Table 26: Millet model specification test (linktest)

GEE population-averaged model		Number of obs =	660
Group variable:	id2	Number of groups =	230
Link:	log	Obs per group:	
Family:	Poisson	min =	1
Correlation:	exchangeable	avg =	2.9
		max =	3
		Wald chi2(2) =	180.42
Scale parameter:	1	Prob > chi2 =	0.0000

pfl_f	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
hat	.9671669	.1029325	9.40	0.000	.7654229	1.168911
hatsq	-.0379002	.082669	-0.46	0.647	-.1999284	.1241281
_cons	.0145126	.0684192	0.21	0.832	-.1195867	.1486118

This test shows that the predicted millet losses (hat) are linearly related to the observed millet losses with the selected statistical model and the included predictors, the link is not different to one (linear test β_hat=1, p=0.749). The lack of fit represented by a nonlinear trend of the predicted millet loss (hatsq) is not significantly different from zero (β_hatsq=0,p=0.647). This means that the selected model is a good millet loss predictor.

Base model for maize

For maize, two outliers with >20% maize loss were excluded. The percent loss showed similar values for the overall means of the three visits (not significant differences over time), the term was excluded from the model.

Harvesting and handling procedures related to maize losses (table 27) are: threshing with modern machines, with a significant contribution of 41.5% increase in the percent loss (p=0.072), and drying time with an increase of 3.6% per additional day (p<0.001). No drying and raised racks have no significant contributions. The stored amount and the structure capacity represent reduction trends of 5.9% by stored ton, and 3.4% by ton of structure capacity (p=0.068 and p=0.094 respectively).

Table 27: Model estimated incidence rate ratios (IRR=e^β) for maize losses

pfl_f	IRR	Std. Err.	z	P>|z|	[95% Conf.	Interval]
treshing						
Machine moderne	1.414917	.2724891	1.80	0.072	.9700715	2.063754
dryingtime	1.036389	.0082829	4.47	0.000	1.020281	1.052751
1.nodrying	1.924125	.9319041	1.35	0.177	.7446878	4.971557
1.raised_r	.5748107	.2339521	-1.36	0.174	.2588688	1.276351
sto_amount_t	.9412188	.0311921	-1.83	0.068	.8820266	1.004383
stru_cap_t	.9661288	.019855	-1.68	0.094	.927987	1.005838
cont3						
Jute bags	2.258269	.5575444	3.30	0.001	1.391946	3.663777
Polypropylene (PP) bags	1.942657	.3731011	3.46	0.001	1.333266	2.83058
Loose on the ground	1.55548	.3049519	2.25	0.024	1.05922	2.284244
other	1.698768	1.099607	0.82	0.413	.4776983	6.041076
pest_c3						
Sun drying	2.100262	.4654299	3.35	0.001	1.360324	3.242683
Pesticide/Insecticide	1.235478	.144761	1.80	0.071	.9819738	1.554427
Others	1.323507	.2220261	1.67	0.095	.9526471	1.83874
inf_insects						
Oui	1.598024	.1636151	4.58	0.000	1.307471	1.953144
inf_rodents						
Oui	1.163215	.1638953	1.07	0.283	.8825249	1.53318
loginc	1.173352	.0722107	2.60	0.009	1.040024	1.323772
1.credit_EAC	.7717671	.1270621	-1.57	0.116	.5589156	1.065679
mean_t2m	.9944558	.0301353	-0.18	0.854	.9371114	1.055309
max_tp2	.0272114	.0655432	-1.50	0.135	.0002424	3.054984
c.mean_t2m#c.max_tp2	1.127123	.0876793	1.54	0.124	.967734	1.312764
cer						
BAROUELI	1.062446	.287845	0.22	0.823	.624732	1.806842
BLA	1.274767	.2746702	1.13	0.260	.8356509	1.944628
SIKASSO	2.245278	.4698424	3.87	0.000	1.489874	3.383692
_cons	.0336683	.0408339	-2.80	0.005	.0031251	.3627296

The reported containers used for storage showed different levels of maize losses, Jute bags 2.26 times higher, Polypropylene bags (PP) 1.94 times higher, and “Loose on the ground” with 55.5% increment compared to “On cobs”. The estimated percent loss for Jute bags is 1.9%, PP bags with 1.6% compared to 0.8% for “On cobs”, as shown in figure 8.

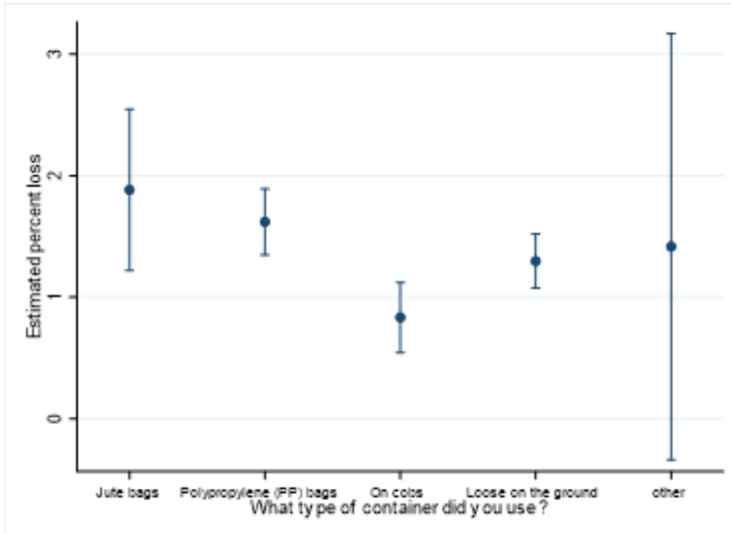

Figure 8: Estimated percentage loss of maize by type of container

For pest control procedures, sun drying showed maize losses 2.1 times higher (p=0.001), the use of Pesticide/Insecticide show an increment of 23.5% in losses (p=0.071), and other pest controls show losses 32.4% higher in comparison to no pest control procedures. Sun drying showed an estimate of 2.9% of maize losses, the use of Pesticide/Insecticide with 1.5%, and others with 1.6% losses in comparison to 1.2% with no controls (figure 9).

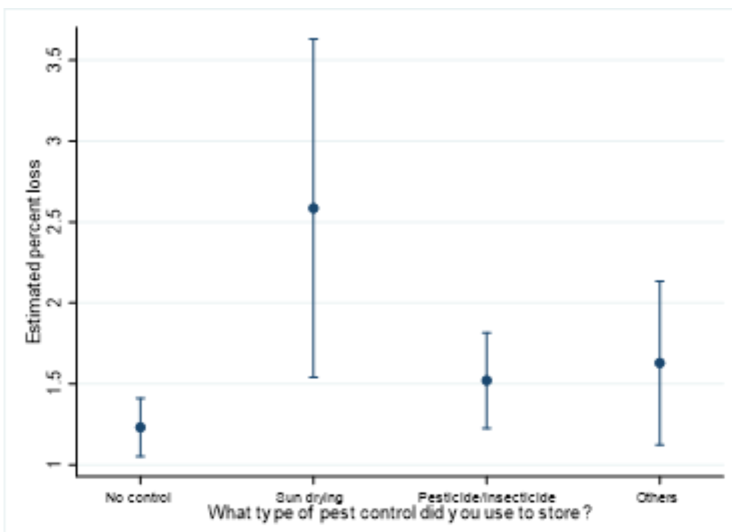

Figure 9: Estimated percentage loss of maize by type of pest control used

Insect infestation showed a 59.8% increment in maize losses (p<0.001) and rodents' infestation does not show a clear trend.

The logarithm of income reported is associated with a 17% increment in loses, and the access to credits reported in the EAC survey shows a small reduction trend of 23%.

The effect of temperature and precipitation shows an interaction effect on maize losses (p=0.124). This can be explained using a contour plot shown in figure 10.

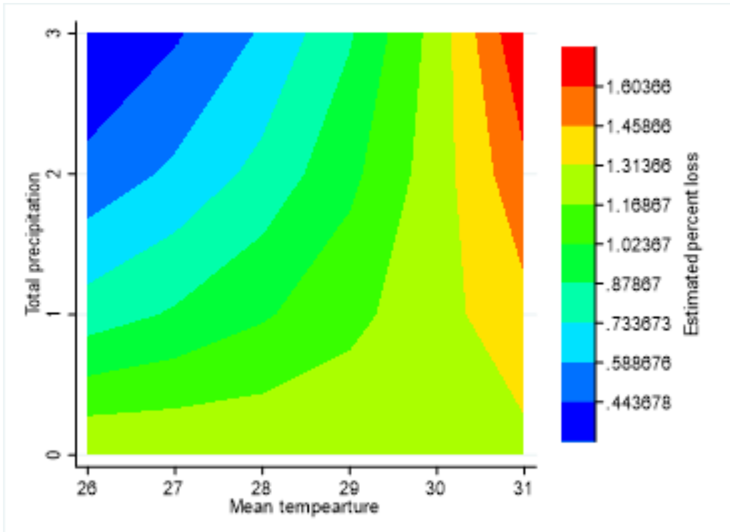

Figure 10: Weather factors interaction for maize loss

The geographic division in Cercles, shows that maize losses are higher in SIKASSO with 2.24% compared to KOUTIALA with 1%, BAROUELI with 1.06%, and BLA with 1.27%, as shown in figure 11.

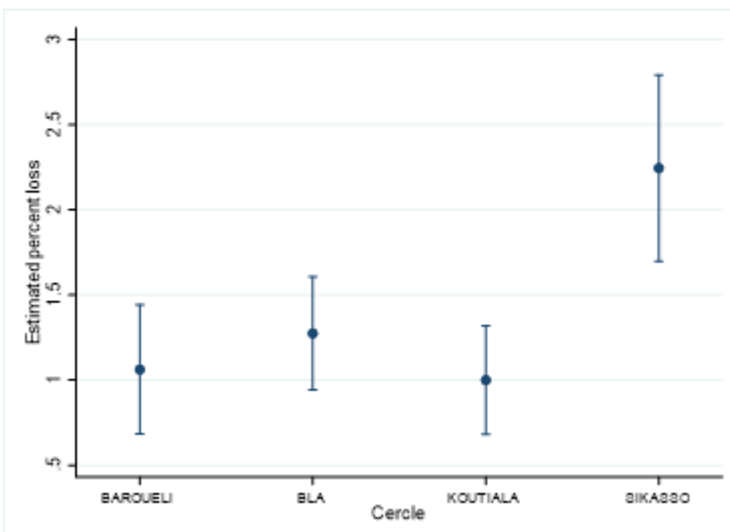

Figure 11: Estimated percentage loss of maize by district

The specification test shows a good linear relationship between the predicted and the observed food loss, model is a good predictor of millet losses. As shown in table 28.

Table 28: Maize Model Specification test (linktest)

GEE population-averaged model		Number of obs =	357
Group variable:	id2	Number of groups =	132
Link:	log	Obs per group:	
Family:	negative binomial(k=1)	min =	1
Correlation:	exchangeable	avg =	2.7
		max =	3
		Wald chi2(2) =	67.82
Scale parameter:	1	Prob > chi2 =	0.0000

pfl_f	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
hat	1.002372	.1347528	7.44	0.000	.7382614	1.266482
hatsq	-.0605397	.1324428	-0.46	0.648	-.3201228	.1990435
_cons	.1017885	.1069361	0.95	0.341	-.1078024	.3113795

The predicted maize loss (hat) has a linear correspondence; it is not significantly different to one (linear test of β_hat=1,p=0.986). The lack of fit represented by (hatsq) is not significantly different from zero (β_hatsq=0, p=0.648). This means that the selected model is a good maize loss predictor.

ANNEX 3

Integrated models using Mali 2023+EAC survey predictors.

Integrated models based on Mali 2023 predictors, EAC predictors, and weather variables are presented. Initial Stata output of all models are omitted, estimated coefficients expressed as rate ratios (RR) are presented and interpreted as a multiplicative contribution to the percent loss or as a percentage of contribution to the percent loss are presented.

Integrated model for millet

The relational structure of food loss drivers to the estimated percent of millet loss, is presented in table 29.

Table 29: Model estimated rate ratios (IRR=e^β) for millet losses

pfl_f	IRR	Std. Err.	z	P>|z|	[95% Conf.	Interval]
S7Q10_APPUI_CONSEIL_TIC	.5121246	.1058529	-3.24	0.001	.3415369	.7679158
S7Q18M1_HERBES_LDP2	1.96514	.4291961	3.09	0.002	1.280817	3.015088
S7Q18M6_INSECUR_LDP2	.2415306	.0949786	-3.61	0.000	.1117503	.5220301
dmechown	.4515556	.1376055	-2.61	0.009	.2484957	.8205471
dmechrent	.459106	.2009309	-1.78	0.075	.1947058	1.082548
q15trloss	1.015604	.007783	2.02	0.043	1.000463	1.030973
q21drloss	1.303166	.3341232	1.03	0.302	.7884193	2.153983
herbicide_qty	1.107086	.0446001	2.53	0.012	1.023034	1.198045
stru_height						
1+ meter	1.793402	.6327382	1.66	0.098	.8981762	3.580914
1.floordry	2.804156	1.221057	2.37	0.018	1.194402	6.583457
temp_wrkrs	.9284314	.0395618	-1.74	0.081	.8540413	1.009301
visit						
2	.1213826	.0783269	-3.27	0.001	.0342677	.42996
3	.0207592	.0217717	-3.69	0.000	.0026576	.1621534
weeksSTO	1.004491	.0161904	0.28	0.781	.9732539	1.03673
age	1.011106	.0069925	1.60	0.110	.9974932	1.024904
mean_t2m	2.184633	.4308195	3.96	0.000	1.484286	3.215431
sum_tp2	34.23924	71.68536	1.69	0.091	.5654533	2073.249
c.mean_t2m#c.sum_tp2	.8950848	.0607855	-1.63	0.103	.7835358	1.022515
_cons	1.82e-10	9.72e-10	-4.20	0.000	5.13e-15	6.45e-06

The first part of the model includes predictor variables from the EAC survey. If the respondent received advisory support through at least one ICT, there is a 49% reduction and if the respondent had experienced insecurity there is a 76% reduction in the expected loss. If encountered difficulties with weeds there is an increase of 96% in expected loss. The use of owned machinery and rented machinery are related to 55% and 54% reduction respectively. An increase in the quantity lost during threshing implies an increase of 1.6% loss, and an increase in the quantity of herbicide used prior to the harvest implies a 10% increase in loss.

The second part of the model includes predictor variables from the Mali survey. If the height of the structure is 1 metre and over, there is a 79% increase in losses, floor drying increases expected millet losses almost 3-fold. An increase in the number of temporary workers is related to a 7% reduction. There is a very significant reduction in millet losses on the second (88%) and third (98%) visit compared to the first visit. The time that the grains have been stored and the age of the household head apparently are not related to millet losses.

The last part of the model includes the weather variables, where the mean temperature and the total precipitation are positively related to millet losses and with a low significant interaction as shown in figure 12.

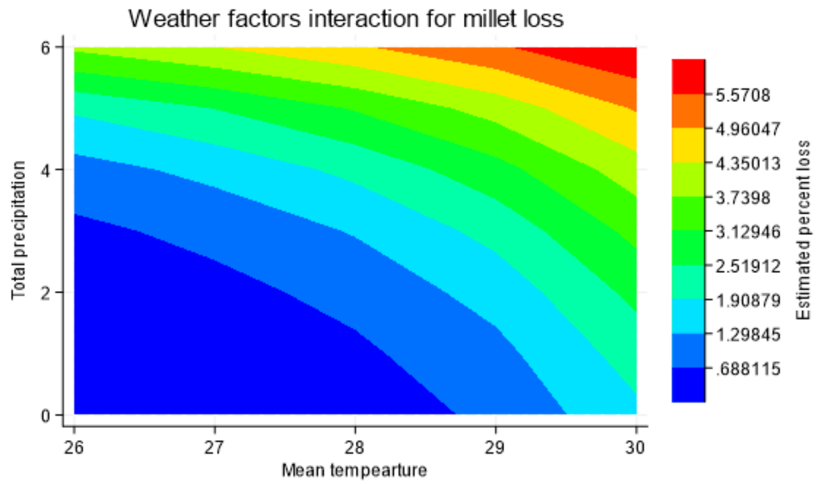

Figure 12: Weather factors interaction for millet losses

The predicted millet loss (hat) has a linear correspondence; it is not significantly different to one (linear test of β_hat=1, p=0.242). The lack of fit represented by (hatsq) is not significantly different from zero (β_hatsq=0, p=0.648). This means that the selected model is a good millet loss predictor as shown in Table 30.

Table 30: Linktest for the integrated model for the percent loss of millet

GEE population-averaged model		Number of obs =	417
Group variable:	id2	Number of groups =	144
Link:	log	Obs per group:	
Family:	Poisson	min =	1
Correlation:	exchangeable	avg =	2.9
		max =	3
		Wald chi2(2) =	147.52
Scale parameter:	1	Prob > chi2 =	0.0000

pfl_f	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
hat	1.078314	.1080408	9.98	0.000	.8665582	1.29007
hatsq	.0885444	.0757308	1.17	0.242	-.0598852	.236974
_cons	-.059412	.1008044	-0.59	0.556	-.2569849	.138161

Integrated model for maize

The relational structure of food loss drivers to the estimated percent of maize loss, is presented in table 31.

Table 31: Model estimated rate ratios (IRR=e^β) for maize losses

pfl_f	IRR	Std. Err.	z	P>|z|	[95% Conf.	Interval]
S7Q10_APPUI_CONSEIL_TIC	.8255416	.1425203	-1.11	0.267	.5885573	1.157948
credit_EAC	.5764279	.1075777	-2.95	0.003	.3998405	.8310041
dmechown	1.124909	.2165826	0.61	0.541	.7713168	1.640598
dmechrent	.7001685	.2015019	-1.24	0.216	.3983224	1.230751
q15trloss	.4449492	.1131849	-3.18	0.001	.2702616	.7325489
herbicide_qty	.9432561	.0350853	-1.57	0.116	.876937	1.014591
inf_insects						
Oui	1.359513	.133997	3.12	0.002	1.120693	1.649225
inf_rodents						
Oui	1.498048	.2085226	2.90	0.004	1.14036	1.967931
1.raised_r	.3254841	.1887154	-1.94	0.053	.1044729	1.014041
cont3						
Polypropylene (PP) bags	.4363329	.0967888	-3.74	0.000	.2824891	.6739602
On cobs	.3470929	.0926495	-3.96	0.000	.2057001	.5856754
Loose on the ground	.5216977	.1123085	-3.02	0.003	.3421184	.7955388
other	.3836276	.2284945	-1.61	0.108	.1193771	1.232817
visit						
2	.4960629	.182632	-1.90	0.057	.241077	1.020746
3	.3400365	.1799753	-2.04	0.042	.1205035	.9595143
weeksSTO	1.008251	.0100363	0.83	0.409	.9887705	1.028115
mean_t2m	1.181673	.1317403	1.50	0.134	.9497302	1.470262
max_tp2	1.11659	.077625	1.59	0.113	.9743577	1.279585
_cons	.0514053	.1565939	-0.97	0.330	.0001312	20.13666

The first part on the predictor variables from the EAC survey, showed a 17.5% reduction in losses, if the respondent received advisory support through at least one ICT; if the respondent had access to credit there is a 42.4% reduction in the expected loss. The use of owned machinery and rented machinery are not clearly related to maize losses. An increase in the quantity lost during threshing implies a reduction of 65.5% in losses.

The second part of the model includes predictor variables from the Mali survey. If they reported insects' infestation and rodent attacks, there is a 36% and a 50% increase in losses respectively. If they use raised racks, there is a reduction of 68.4%. The reported containers used for storage showed different levels of maize losses, Polypropylene bags (PP) 56.4% less losses, On cobs 65.3 and "Loose on the ground" with 47.8% less compared to "Jute bags". The estimated percent loss for Jute bags is 2.6%, PP bags 1.1%, Loose on the ground with 1.4%, and 0.9% for "On cobs", as shown in figure 13.

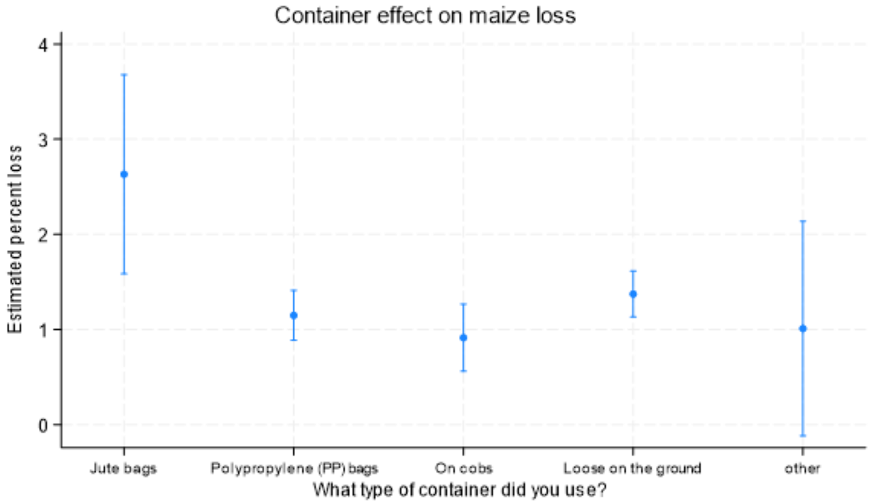

Figure 13: Estimated percentage losses of millet by type of container

There is a very significant reduction in millet losses on the second (50.4%) and third (66%) visit compared to the first visit.

The last part of the model includes the weather variables, where the mean temperature and the maximum precipitation seem positively related to maize losses but with no significance.

Table 32: Link test for the integrated model for the percent loss of maize

GEE population-averaged model		Number of obs =	289
Group variable:	id2	Number of groups =	105
Link:	log	Obs per group:	
Family:	negative binomial(k=1)	min =	1
Correlation:	exchangeable	avg =	2.8
		max =	3
		Wald chi2(2) =	30.83
Scale parameter:	1	Prob > chi2 =	0.0000

pfl_f	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
hat	.8937476	.1680328	5.32	0.000	.5644093	1.223086
hatsq	-.0208961	.2070867	-0.10	0.920	-.4267786	.3849863
_cons	.1197221	.1229901	0.97	0.330	-.1213341	.3607782

The predicted maize loss (hat) has a linear correspondence; it is not significantly different to one (linear test of β_hat=1,p=0.527). The lack of fit represented by (hatsq) is not significantly different from zero (β_hatsq=0, p=0.92). This means that the selected model is a good maize loss predictor.