Centre of Excellence in Agri-Food Systems and Nutrition

CE-AFSN | Working Paper #1 | Nov 2025

Toward Synergy in Mozambique's Agricultural Production Statistics: Complementarity between Integrated Agricultural Survey and Statistics from Space Approach

Lourenço Manuel, Jackelya Araujo da Silva, Rolf de By and Rui Benfica

Ministry of Agriculture. Food and Rural Affairs
Republic of Korea

November, 2025

Toward Synergy in Mozambique's Agricultural Production Statistics: Complementarity Between Integrated Agricultural Survey and Statistic from Space approach

Lourenco Manuel, Jackelya Araujo da Silva, Rolf de By and Rui Benfica

Abstract

Agricultural production data plays a critical role in economic development, especially for developing countries like Mozambique, as it informs more effective policymaking on agricultural investments, subsidies, and initiatives at local, national, and regional levels. Currently the Integrated Agricultural Survey ("Inquérito Agrário Integrado" - IAI) serves as the official source of agricultural information in Mozambique. It collects various types of data at farmers level, including estimates of crop cultivated areas, which is based on farmers self-reporting and the measurement of two farmers' plot sizes in each enumeration area (EA). Although IAI employs rigorous methodologies, the accuracy of cultivated area estimates may be limited due to low literacy levels among farmers and the prevalence of intercropping production systems, which complicate the determination of the area occupied by each crop. In recent years, digital technologies and satellite remote-sensing methodologies have emerged as promising tools for estimating crop areas with greater precision and timeliness. In this paper, we advocate for the application of such methodologies – referred to here as "Statistic from Space (SFS)" – for crop area estimation. We use two data sources: IAI 2023, representing official government data collected by the Ministry of Agriculture in collaboration with the National Institute of Statistic (INE); and SFS data collected in 2025 under the "Statistics from Space" project. We conduct a simulation study using both sources of data and evaluate the performance of each approach using statistical metrics such as the range of confidence interval, margin of error, and the asymptotic relative efficiency of the SFS approach compared to its counterpart, the IAI. The results indicate that the SFS proposed in this paper outperforms the IAI across all performance measures. These findings suggest substantial complementarity and synergy between the SFS approach and IAI methodologies. We conclude that integrating the SFS approach into the planning and implementation of the IAI could significantly enhance the precision of the crop area estimates in Mozambique.

Key words: *statistics from space; integrated agricultural survey; complementarity; crop areas estimation.*

1. Introduction

Agricultural production data is a vital resource for enhancing agricultural productivity and ensuring food security in developing countries such as Mozambique. Among its main roles, agricultural data supports evidence-based policymaking on agricultural investments, subsidies and initiatives at both community and national levels.

In Mozambique, the official source of agricultural production statistics was established in 1993 with the launch of the "Trabalho de Inquérito Agrícola (TIA)", whose first edition also took place that year. TIA was designed to collect comprehensive annual data on crop and livestock production to meet the data needs of the Ministry of Agriculture. Subsequent TIA surveys were implemented in 1996, 2002, 2004, and 2008. In 2012, the Ministry of Agriculture introduced the "Inquérito Agrário Integrado (IAI)" which effectively merged TIA with the "Sistema de Aviso Prévio" (Early Warning System), creating a more integrated and efficient data framework.

Currently, the IAI remains the primary source of agricultural official data in Mozambique. It intended to be conducted annually to provide essential information on inputs and outputs of agricultural activities for policy planning and performance evaluation of the agriculture sector. However, in the last five years, the survey was only implemented in 2020 and 2023 (MADER, 2024)

The Statistics from Space project (SFS) offers a complementary approach to agricultural data collection, with a particular focus on estimating crop areas at the provincial and district levels estimates. The SFS approach leverages satellite remote-sensing data and artificial intelligence augmented analytics to produce and disseminate accurate crop area statistics in a timely manner.

Funded by the Government of the Republic of Korea through its Ministry of Agriculture, Food and Rural Affairs (MAFRA), the SFS project aims to support Mozambique's Ministry of Agriculture by providing timely production estimates for major crops across three provinces: Gaza, Manica and Zambezia. The goal is to ensure that all stakeholders in the agricultural chain can access and use data for decision-making.

The SFS was implemented by four lead institutions, each responsible for a distinct component of the project: The Centre of Excellence in Agri-Food Systems and Nutrition at Eduardo Mondlane

University (CE-AFSN UEM) led the ground-truthing and digital data collection component; the International Food Policy Research Institute (IFPRI) oversaw the stakeholder's engagement component; the ITC at University of Twente (Netherlands) led the component of Enhanced area sampling frame; and the Seoul National University (SNU) from South Korea led the Analytical framework component.

In this paper, we explore the potential for complementarity between the IAI and the SFS methodological approaches. Using data from IAI 2023 and SFS 2025, we conduct comparative analyses to address the persistent challenge of estimating crop areas, particularly in intercropping production systems which mostly characterizes the Mozambique's agricultural landscape. Traditional IAI methods rely heavily on farmers self-reported data, which may not be accurate due to their low literacy levels. Although IAI methodology randomly selects 2 households in each of the Enumeration Areas to effectively measure the plot size of these farmers using a GPS device, accurately estimating the area occupied by each crop remains difficult in intercropped plots. The SFS approach addresses this challenge by incorporating digital technologies and satellite remote-sensing to estimate crop cultivated areas with greater precision and in a timely manner. This paper evaluates the performance of both methodologies and highlights the potential benefits of integrating SFS into the IAI framework.

2. METHODOLOGY

This study utilizes two primary data sources: the Integrated Agricultural Survey (IAI 2023) and the Statistics from Space (SFS 2025). The former corresponds to the government official statistics collected by the Ministry of Agriculture in collaboration with the National Institute of Statistics (INE). The second was collected under the Statistics from Space project, implemented in Mozambique by four institutions namely International Food Policy Research Institute (IFPRI), Seoul National University (SNU), Faculty of Geoinformation Science and Earth Observation of University of Twente (ITC) and the Centre of Excellence in Agri-Food Systems and Nutrition (CE-AFSN). The following subsections depict the detailed methodology used in both approaches (IAI 2023 and SFS).

2.1 Integrated Agricultural Survey (IAI 2023)

The IAI 2023 sampling approach was based on the Master Sample Frame for Agricultural Surveys generated from Mozambique's fourth General Population and Housing Census carried out in 2017. The sampling frame applies a probabilistic approach in two stages. In the first step the Primary Sampling Units (PSU) also known as Enumeration Areas (EA) were randomly selected within each Province and District, considered as strata. A total of 2366 EA were selected, excluding 9 districts of Cabo Delgado Province due to instability. However, of these, 39 EAs were not surveyed due to accessibility challenges and depopulation, resulting in an attrition rate below 2%, which does not compromise sample representativeness. A listing of households was also carried out in each EA including one question to identify the medium-scale farms, which were included in the sample with probability equal 1 in the second sampling stage. Secondly, within the randomly selected EA a total of 11 Households classified as smallholder famers in which 8 were considered as "the main sample" and the remaining 3 were used for replacements and were selected using a random sampling systematic approach. Still in the selected EA a census of all medium-scale famers was conducted. For the large-scale farmers, a census was intensively also carried out in all districts of the country. The Integrated Agricultural Survey 2023 used the same sample size determined in IAI 2020 which corresponds to around 25 thousand households (small and medium-scale farmers) across 145 districts. The Integrated Agricultural survey applies a structured questionnaire to collect data among randomly selected small, medium, and large-scale farmers. The survey aims to collect several data such agriculture and livestock production, sales, household income, agricultural production system and food security indicators (MADER, 2024).

2.1.1 Agricultural Production Estimates

Crop area estimates

Crop area cultivated, yield and production are key variables collected in agricultural surveys. The area and crop production are the main variables used to determine crop yield, which is a measure of productivity. In each EA, two (2) out of eight (8) households are randomly selected for direct measurements using a GPS device and visual inspection. The enumerator keeps the GPS device "on" while he walks along the entire perimeter of the plot from a specific starting point and at the end of the process the enumerator records the plot area. For the remaining households, the plot sizes are recorded based on the farmer's self-reports. However, it is important to point out that the majority of the smallholder farmers practice intercropping system which is a challenge for estimating crop-specific areas.

Crop production estimates

Production estimates were based on farmer recall. Right after the harvest, the interviewers visit farmers to estimate production. Interviewers use a sheet to record the production (of each required crop), ask the farmer how much he/she harvested. For crops that are harvested at a specific time (like beans or corn), it is easy to remember for farmers. For crops that are harvested over extended periods (e.g., cassava and sweet potato), this is a little more difficult. Even for these crops, they are usually harvested for 2 or 3 months during the season. Enumerators are trained to ask farmers how much they harvested each week during that 2-3months period and then multiply by the number of months to get the season's output.

Another aspect related to the difficulty of measuring production is the use of non-standardized unit measures that farmers use to determine their production. Several studies have been carried out to determine the weights of many of these non-standard units, so that the conversion factors are available and known for subsequent conversion of production to a standard unit such as kilogram.

Crop yield estimates

Estimating crop yields is inherently challenging, particularly in smallholder farming systems where intercropping is common (Murphy et al. 1991). In the IAI 2023, harvest yield was measured through direct interviews with the household using the farmer recall method. Farmers are interviewed after harvest operations and are requested to recall the produced quantity of each crop per plot. This method post-harvest yield estimation is usually carried out at the farmer's house or the place where the crop is stored, allowing surveyors to cross-check estimates against available storage capacity (Casley and Kumar, 1988). The method has the potential to provide accurate agriculture production estimates in countries that have achieved higher levels of mechanization, commercialization and record keeping. It is useful when farmers are literate and with some knowledge. Another procedure used for yield estimation consists on the farmer's forecasts, where the farmer is asked to estimate the expected amount of production amount in a plot that has not yet been harvested. The method is useful when used to forecast agricultural production 15 to 20 days before harvest. The method is useful when farmers are literate and connoisseurs (Casley and Kumar, 1988).

For the IAI 2023, the farmer recall method was used to estimate crop yields, relying on previously collected data on crop area cultivated and production estimates across various crops practiced by the households.

2.2 Statistics from Space Approach (SFS)

Under the proposed methodological approach, our objective was to quantify agricultural activity across three provinces, Gaza, Manica and Zambezia, by estimating the extent of land devoted to a variety of crops, which leads to an area size estimate (in hectares) for each of these crops. The method proposed here is spatially explicit, and this allow us to determine estimates of crop area size for any geographic entity (for instance, district) within the provinces.

The approach makes use of stratified sampling design. In this context, a *stratum* is defined as a maximal grouping of spatial patches within which we expect annual agricultural production to be relatively homogenous. By this, we mean to state that the parameters of primary input to crop farming are quite similar. These parameters include soil and elevation characteristics, weather conditions and level of farming intensity. Two locations are considered part of the same stratum when such parameters are sufficiently similar, ensuring that the sampling units within each stratum are comparable in terms of agricultural potential and practices.

The process of stratum derivation is outlined below. Once strata have been defined, we identify *sampling areas* (SAs) as fieldwork locations where crop data is collected. A financially sustainable number is 100 SAs per province; we hold 50 SAs in reserve per province. Typically, multiple SAs are selected within a stratum and together they serve to represent the stratum and allow derivation of crop statistics, such as cultivated area per crop. The basis for this is the *crop cover percentage* $ccp(c, s)$ of a crop c in a stratum s. It is determined from the collected fieldwork data. Under the assumption of production homogeneity, production is similar throughout the stratum's extent.

This set-up enables the derivation of crop statistics for any geography g of interest. The process begins by identifying which strata intersect with g, and calculating the area sizes of each intersection. For a given crop c, we multiply each area size by the ccp of the associated stratum, and the resulting values are summed up across all relevant strata to obtain the total estimated area size of crop c within geography g.

2.2.1 Stratum Derivation

To derive meaningful strata spatially, we characterize the underlying environmental and agronomic conditions, especially soil type, elevation characteristics, weather conditions and level of farming intensity. This spatial characterization enabled the identification of about 20 strata for each province. These derived strata are designed to remain stable across multiple agricultural seasons or campaigns, so thereby allowing the detection of crop production trends and facilitating interannual comparisons. The following base data sets were used to derive the strata:

1. *Land forms*: For the three provinces together, there are nine land forms, including surface freshwater. This is one of the data layers served by the SFS data platform[1]. The data was obtained from Cenacarta[2], Maputo and was produced in 2012.
2. *Farming intensities*: Four different classes were identified, regular, scarce, irrigated, and other, the latter interpreted as 'no farming'. This data for Gaza also was obtained from Cenacarta, Maputo and was produced in 2012. For Manica and Zambezia, our classification derives from the crop/no-crop maps produced during the project by ITC team[3].
3. *NDVI profiles*: This is a country-wide raster at 1×1 km resolution that labels each cell from a label set with 61 labels. Such a label characterizes the cell's NDVI behaviour over the analyzed period of 20 years, and typically associates with a sinusoidal form of NDVI development across a year's dekades. This means that cells within the same land category

[1] https://mapfra.itc.utwente.nl/
[2] Mozambique National Cartography and Remote Sensing Centre (CENACARTA)

[3] Crop map for major crops in Mozambique. www. mapfra.itc.utwente.nl

with equal label display similar NDVI behaviour. We consider it a good single proxy for a combination of biophysical vegetation growth conditions.

A total of 61 strata (27 in Gaza, 17 in Manica and 17 in Zambezia province) were defined and within each stratum a set of *sampling areas* (SAs) were identified as fieldwork locations where crop data was collected. A total of 303 SAs were defined as representative for the 3 provinces, i.e, 95 SA in Gaza, 100 SA in Manica and 108 SA in Zambezia province. The approach also included the identification of additional 50 SAs per province to be used for replacements.

Each SA corresponds to units of 500m x 500m blocks and within each SA there are several segments of cultivated and non-cultivated areas (Figure 1). The 303 sampling areas of 500m x 500m were distributed across several districts in Gaza, Manica and Zambezia provinces, which constitutes the primary sample used for data collection. In addition, a list of replacement SAs was provided as backup in cases of inaccessibility of some sites from the primary sample. Table 1 presents the sample size distribution (number of SA) in each district within the three provinces.

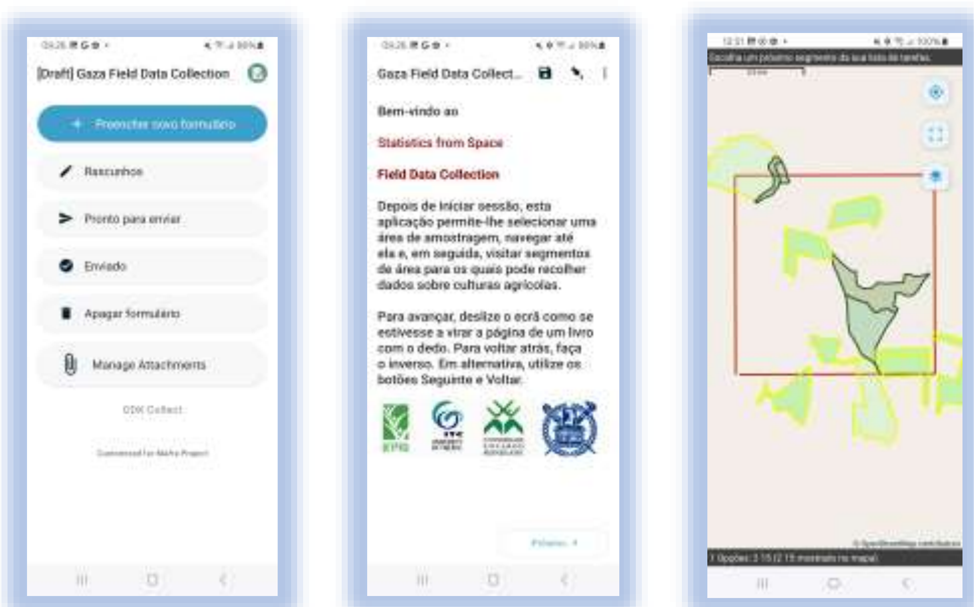

Figure 1. Data collection tool showing one sampling area exhibiting several segments.

Table 1. Sample size distribution per province and district

Zambezia		Gaza		Manica	
Districts	# of SAs	Districts	# of SAs	Districts	# of SAs
Milange	16	Chokwe	12	Barue	10
Lugela	1	Chibuto	14	Guro	13
Gurue	10	Chicualacuala	5	Tambara	1
Mocubela	6	Chigubo	5	Gondola	10
Ile	4	Mapai	7	Manica	12
Mocuba	7	Massangena	3	Vanduzi	3
Maganja	2	Guija	7	Macate	4
Alto Molocue	8	Mabalane	6	Mussorize	16
Morrumbala	12	Massingir	5	Machaze	13
Derre	2	Mandlakazi	5	Sussundenga	18
Molumbo	15	Xai-Xai	3	**Total**	**100**
Mulevala	6	Chongoene	2		
Namacura	3	Limpopo	9		
Namaroi	3	Bilene	12		
Nicoadala	5	**Total**	**95**		
Pebane	6				
Chinde	1				
Quelimane	1				
Total	**108**				

2.2.2 Data collection

Data collection was carried out in three distinct time periodsts across the target provinces: from 14th February to 01st March 2025 in Manica, from 08th to 22nd March 2025 in Zambezia, and from 14th to 28th April 2025 in Gaza. Fieldwork was conducted by four teams per province labeled, as "orange, green, black and blue", with each team having three enumerators. The teams were created based on the spatial distribution of the target sites (sampling areas) to be visited within the 3 provinces. They walk a circuit in the sampling area and visit the field segments previously assigned to them. Per segment, they observed/registered land cover, land use, crops grown and their respective ground cover percentages. Further details on above-ground cover were also observed. All data was entered into our information stack with a mobile phone application which exercises a level of early data curation to enhance quality and consistency.

Overall, the data collection process was organized in a way to maximize the efficient use of limited available resources without compromising data quality or the objectives of the assignment. In addition to conduction field observations survey within the selected SA, a structured questionnaire was also administered to farmers with plots in those SA in each district covered by the study.

Each team consisted of one supervisor, three enumerators, and one driver. Supervisors were recruited among lecturers from different units of Eduardo Mondlane University with strong experience in fieldwork and holding at least a Master degree level. Lecturers from the Higher Polytechnic Institute of Manica (ISPM) and University of Zambezia were also part of the supervisors and they led the teams and were responsible for organizing the field-visits in terms of contacting local authorities and pre-site identifications.

The teams worked in close collaboration with agricultural extension agents at district level and received strong support from the community leaders at local level (targeted SA). The community leaders played an important role in providing authorization to the teams to get in the fields as well as by helping them to identify the target SA.

2.2.3 Crop Cover percentage determination

We focus here on how crop cover percentage within a stratum is determined from the fieldwork data. For a given crop, and each visited segment within a stratum, we determine the area purely covered by that crop in m^2. These crop-specific areas are then summated across all segments in the stratum. We also summate the total area of segments visited by fieldwork enumerators, whether segments are cultivated or not. The two summations determine the crop cover ratio for the crop in this stratum.

Once we have determined all cover percentage for all combinations of crop and stratum, we then determine a crop statistic for any crop and any geography with the area of the three provinces. This is done as the summation of the crop cover percentage multiplied by area size of intersection of that geography and the stratum. This will tell us the area size devoted to the crop in the geography. A crop cover percentage for the crop in the geography is easily determined by also involving the geography's area size.

2.3 Comparative analysis of cultivated area estimates between IAI and SFS

The previous sections detailed the methodologies used in each of the approaches under analysis in this study. IAI 2023 dataset corresponds to data referred to the 2022 agricultural season while SFS data was recently collected and corresponds to the 2025 agricultural growing season. To enable a fair comparison between the two approaches, we used historical data of previous IAI campaigns from 2012 to 2023 to forecast cultivated area for major crops and compare the estimates with the SFS data. We apply a trend analysis using ordinal least square (OLS) estimators to predict the crop area estimated for 2025 agricultural season and we denoted it by "IAI 2025". In addition, we selected the most commonly used statistical measures to assess the performance of both approaches (IAI and SFS).

2.3.1 Statistical measures of performance

The comparison between the two methodological approaches (SFS and IAI) was conducted using widely recognized and interpretable statistical metrics. Specifically, we assessed performance using such as the margin of error, the range of the 95% confidence interval, and estimates of asymptotic relative efficiency.

i. Margin of Error

The margin of error is a key statistical metric used to evaluate how accurate the estimates of a certain characteristic from the sample can be used to infer to the characteristics of the population (Gujarati, 2009). For example, it helps evaluate how closely a sample mean approximates the true population mean. A smaller margin of error indicates greater precision and reliability in the estimate. According to Wackerly et al., (2002), the margin of error is given by equation 1:

$$ME = Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, (1)$$

where ME – is the margin of error; $Z_{\alpha/2}$- quantile from the standard normal distribution; $\frac{\sigma}{\sqrt{n}}$ – is the standard error.

ii. Range of the Confidence Interval

The confidence interval (CI) is another well-known statistical metric used to estimate population characteristics through a representative sample (Mood et al., 1974). Basically, the CI provides an interval (a range) of plausible values within which the true population value, e.g., the population mean, will fall, i.e, CI indicates with a certain level of confidence (e.g., 95%) the possible values of the true mean. The range of the interval is the difference between the upper and lower values of the CI. The smaller the range of the CI the better. This means that if two confidence intervals are estimated, one should choose the CI with the smallest range given that that CI estimates the population parameters with more precision. According to Wooldridge (2009), the Confidence Interval and its range (RCI) is given by equation 2:

$$CI(\theta) =]\hat{\theta} - ME; \hat{\theta} + ME[$$

$$RCI = 2ME \text{ , (2)}$$

where CI – is the confidence interval; θ – is the population characteristic (parameter); ME – margin of error; $\hat{\theta}$ – estimator of parameter θ; RCI – range of the confidence interval.

iii. Asymptotic Relative Efficiency (ARE)

According to Wang and Lin (2005), the asymptotic relative efficiency (ARE) of estimator $\hat{\boldsymbol{\theta}}$ is given by the ratio of two variance estimators of θ. In this paper we use the ARE to compare the performance of the SFS methodology vs the IAI approach. The ARE is denoted by:

$$\mathrm{ARE} = \hat{\mathrm{V}}_{\mathrm{sfs}}/\hat{\mathrm{V}}_{\mathrm{IAI}} \text{ ,}$$

where $\hat{V}_{sfs}$ and $\hat{V}_{IAI}$ are the estimated variances of the $\hat{\boldsymbol{\theta}}$ using the SFS and IAI approaches, respectively. The smaller the variance of the estimator the more efficient it is, and therefore the better. That means that if two estimators are used to estimate a population parameter, one should choose the estimator with the lowest variance.

2.3.2 Simulation study for performance evaluation

In order to estimate the statistical performance metrics, namely margin of error, confidence interval range, and asymptotic relative efficiency, a simulation study was conducted using data collected from both the IAI and SFS methodologies. The study employed a resampling approach applying bootstrap methodology with replacements in which several samples were generated for each of the cases (SFS and IAI) considering the same sample size of IAI 2023 and SFS data. During the resampling process the stratified sampling was applied in both cases, that is, for the IAI approach the province and districts were considered as strata and the Primary Sampling Units (PSU/EA – Enumeration Areas) as clusters and the sampling units were the farmers. For the SFS case the farming intensity with four classes (regular, scarce, irrigated and no farming') was used as strata and the PSU/SA (Sampling Areas) as the clusters and the segments as sampling units. The process of simulation was carried out using 100 samples generated for each case (IAI and SFS). In each of the bootstrapped samples generated, the estimates of crop planted area were generated using the IAI and SFS approaches and those estimates recorded. Upon the completion of the simulation, the above statistical metrics were estimated and recorded for both cases (IAI and SFS), enabling a robust comparative analysis of estimator precision, reliability and efficiency.

All simulation procedures and analysis were carried out using software STATA version 17 and R packages.

3. RESULTS AND DISCUSSION

This section presents the results of crop area estimates for the main crops using both methodologies. Firstly, the results start with the trend analysis to forecast cultivated areas of selected crops for 2025 agricultural season using IAI historical data. Lastly, the results of the simulation study to compare both approaches are reported.

3.1 Trend analysis of cultivated area based on IAI data

Figure 2 reports the results of the forecast analysis of cultivated area for four main crops (maize, cassava, beans and small peanuts) using the historical available IAI data from 2012 to 2023. In each of the provinces the adjusted OLS model is used to predict the estimated area for 2025 agricultural campaign per province in the selected crops. These crops are representatives for the groups of cereals (maize), roots and tubers (cassava), beans and oilseed (small peanuts). It is noticed that in all provinces (Manica, Zambezia and Gaza) there is a significant increase of the cultivated area for maize, beans and small peanuts specially in Zambezia where the estimates are higher than the other provinces. For cassava case the analysis shows a slight decrease in crop area estimates for Manica and Gaza while Zambezia province continues increasing the cultivated areas. These results are consistent with MADER (2024).

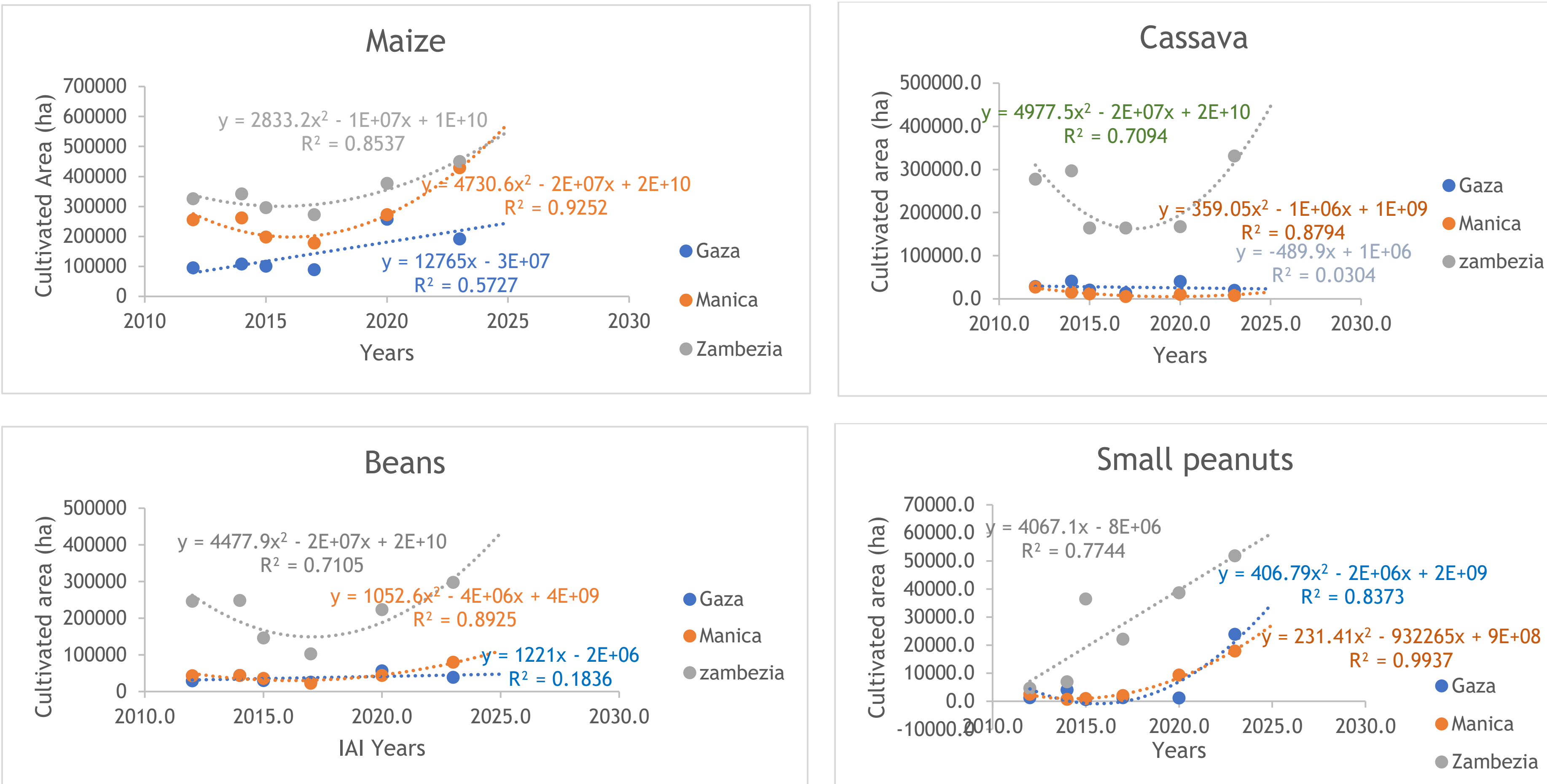

Figure 2. Trend analysis of cultivated area for **maize, cassava, beans and small peanuts** in each province based on IAI data from 2012 to 2023. Cultivated area for 2025 was estimated based on OLS equations.

3.2 Crop Area Estimates

Figures 3 depicts the results of crop cultivated area estimates across the three provinces considering the three data sources: IAI 2023 (observed data), IAI 2025 (forecasted via trend analysis), and SFS 2025 (recently collected data). The results are focused on four major crops, each one representing a major crop group: maize for cereals, cassava for roots and tubers, beans and small peanuts for beans and oilseeds groups. It is noticed that regardless the source of data used to estimate crop areas, maize is the most cultivated crop followed by cassava specially in Zambezia province where crop area estimates were the highest.

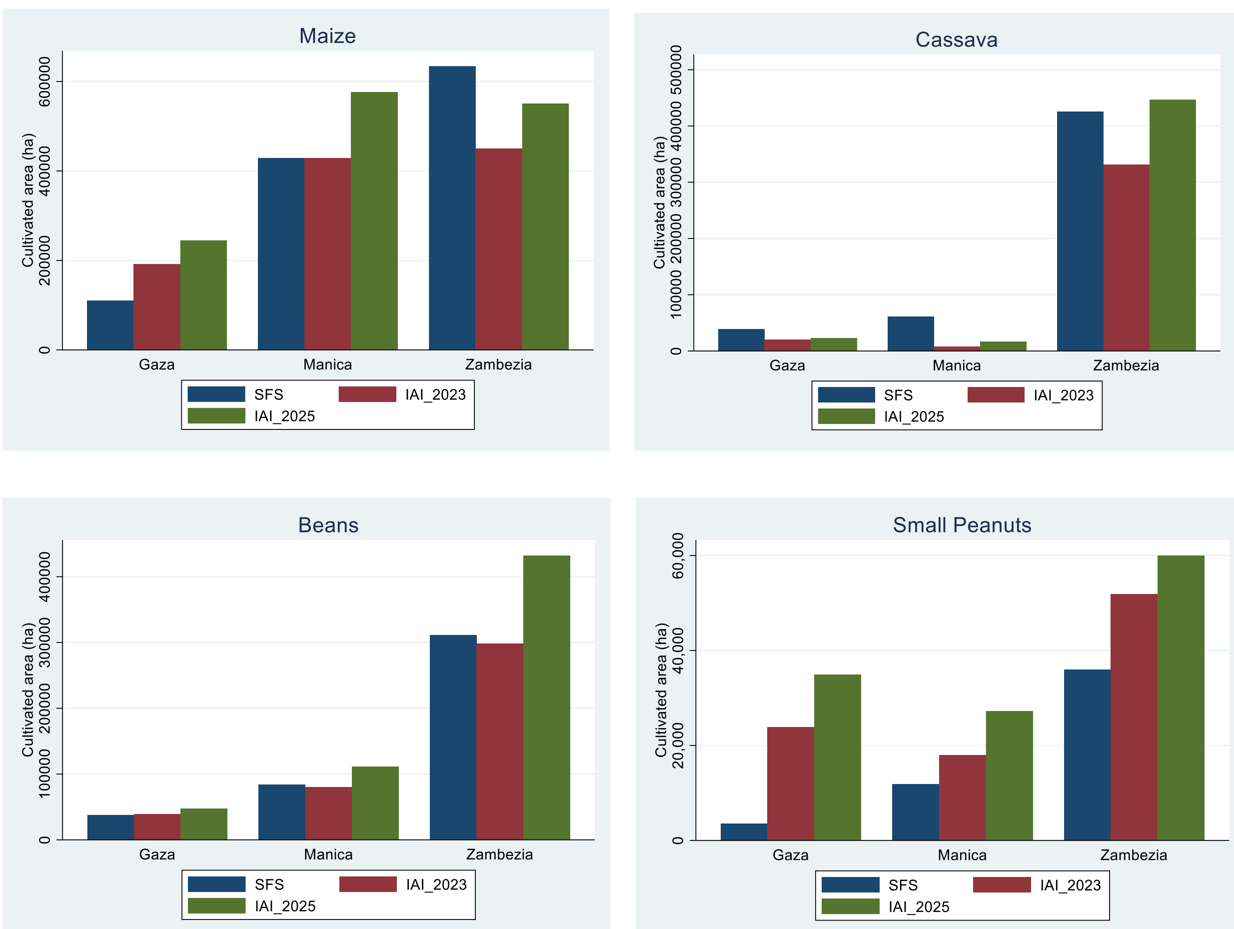

Figure 3. Cultivated area estimated for maize, cassava, beans and ground nut in each province under SFS and IAI data.

In general, crop area estimates for beans and oilseed crops (small peanuts) under SFS approach generates lowers estimates across the three provinces when compared to IAI 2025. This trend is also observed for maize crop with slight variations in Zambezia province where the SFS approach generated higher maize cultivated area than IAI 2025 data. For cassava case, the crop area estimates shows that the trend slightly changes specially for Manica and Gaza province where SFS generated higher estimates although for Zambezia province the SFS approach generated lower cultivated area of cassava.

The results above reported in Figure 3 using both approaches aimed to get a sense of the crop estimated areas under each methodology. It is important to point out that the comparison of both methodology is carried out using appropriate statistical metrics depicted in section 3.

3.3 Comparison of SFS and IAI approaches

This subsection presents results of the simulation study carried out using the SFS and IAI approaches based on the three statistical metrics measures of performance: the range of confidence interval, the margin of error and the asymptotic relative efficiency (ARE) of the SFS estimates compared to its counterparty IAI.

i. **Range of confidence interval**

Table 2 reports the results of the bootstrap 95% confidence interval of the crop cultivated area estimate for maize, cassava, beans and small peanuts across the three provinces using IAI and SFS data. The confidence interval is a statistical technique that provides an idea of the set of possible values of the true mean and a small range of the confidence interval indicates that the true mean was estimated with high precision. In general, the results shows that the ranges of 95% confidence interval of the cultivated area for all crops across the three provinces, obtained using the SFS approach are smaller than the IAI approach (Figure 4). These results suggest substantial gains in precision of cultivated area estimates for all crops when using the SFS approach. The reduced interval widths reflect lower variability and improved reliability of the crop area estimates.

Table 2. Bootstrap 95% Confidence Interval using IAI and SFS data

Crops	**Province**	**IAI_data**				**SFS_data**			
		Mean	**Stand. Error**	**Lower**	**Upper**	**Mean**	**Stand. Error**	**Lower**	**Upper**
Maize	Zambezia	450321.3	25838.3	399679.2	500963.5	621151.0	3877.3	613551.7	628750.4
	Manica	429002.8	32859.66	364599.1	493406.6	392441.5	3231.2	386108.5	398774.5
	Gaza	191671.6	14881.83	162503.7	220839.4	138212.3	1914.2	134460.5	141964.2
Cassava	Zambezia	330722.9	21345.61	288886.3	372559.5	373741.8	4022.1	365858.7	381624.9
	Manica	7725.546	1912.273	3977.559	11473.53	31765.3	268.6	31238.9	32291.8
	Gaza	19548.03	2722.689	14211.66	24884.41	41054.6	484.2	40105.5	42003.7
Beans	Zambezia	25354.56	3381.014	18727.89	31981.22	391358.2	2235.3	386977.2	395739.3
	Manica	25959.71	4585.326	16972.64	34946.78	53716.3	482.3	52771.1	54661.5
	Gaza	7631.232	1307.892	5067.81	10194.65	47500.4	619.8	46285.7	48715.2
Small Peanuts	Zambezia	51824.81	5348.518	41341.91	62307.71	33962.9	902.4	32194.3	35731.6
	Manica	17913.59	3381.692	11285.6	24541.59	6525.3	184.2	6164.4	6886.3
	Gaza	23844.49	3256.804	17461.27	30227.71	4658.8	259.9	4149.5	5168.1

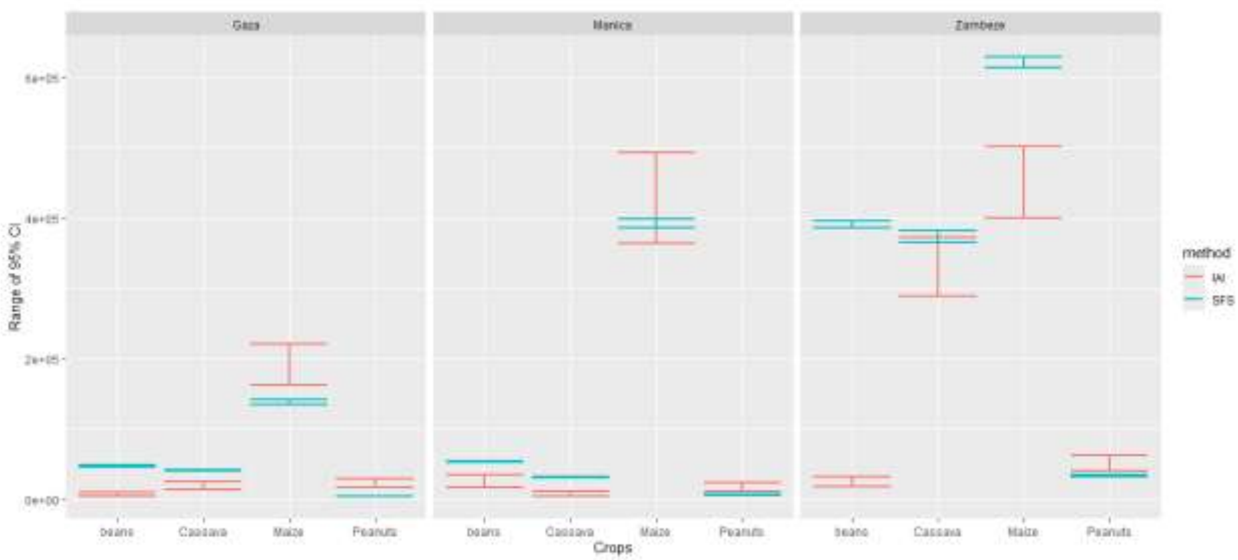

Figure 4. Range of 95% confidence interval of cultivated area for maize, cassava, beans and small peanuts across three provinces using IAI and SFS approaches.

ii. Margin of Error

Figure 5 depicts the estimates of the margin of error for crop areas estimates using the SFS and IAI data for the four crops selected in this study. It is noticed that the highest margin of error estimates are observed under the IAI approach specially for maize crop across the three provinces. In general, likewise the range of the 95% confidence interval, the margin of error estimates obtained using the SFS approach are smaller when compared to its counterpart IAI approach. This suggests that inferences carried out using estimates from the SFS approach introduce lower sampling error. Consequently, inferences drawn from SFS-based estimates are likely to be more accurate and dependable, especially in complex agricultural settings such as intercropping systems prevalent in Mozambique.

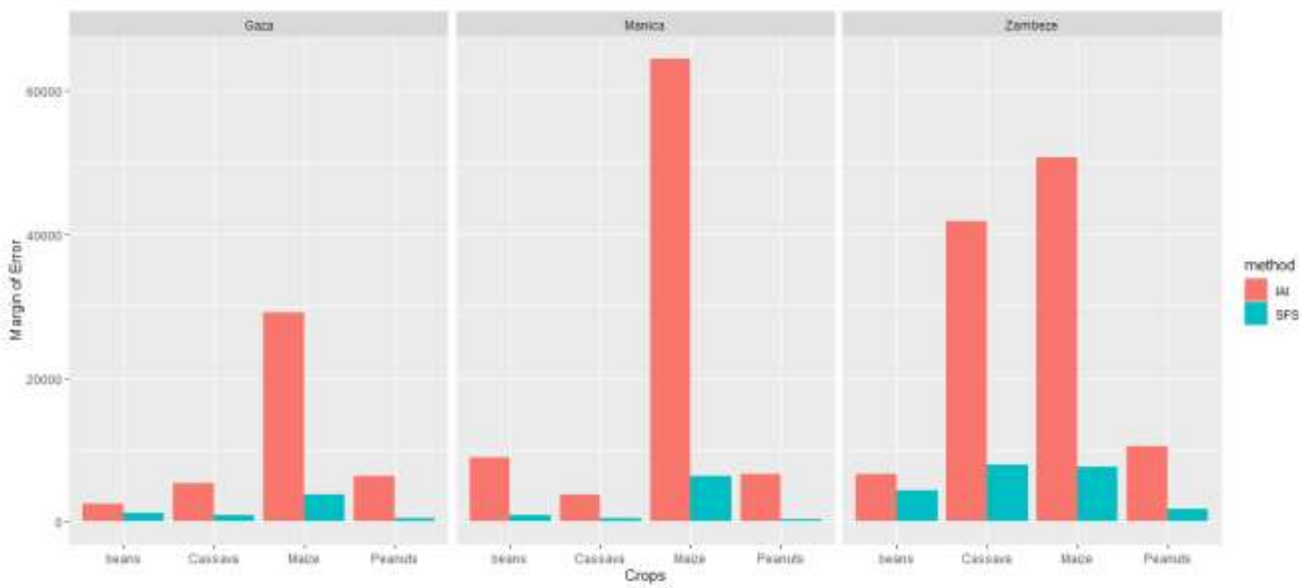

Figure 5. Margin of error estimates for distinct crops under IAI and SFS approaches

iii. Asymptotic Relative Efficiency (ARE)

Figure 6 reports the estimates of asymptotic relative efficiency for the SFS approach relative to its counterpart IAI for the four crops across the three provinces included in this study. The results reveal that estimates of ARE for the SFS approach compared to its counterpart IAI varies between

5% to 65%. It is noticed that the lowest estimates of ARE are observed in small peanuts crop while the highest ARE estimates were obtained in beans (Figure 6).

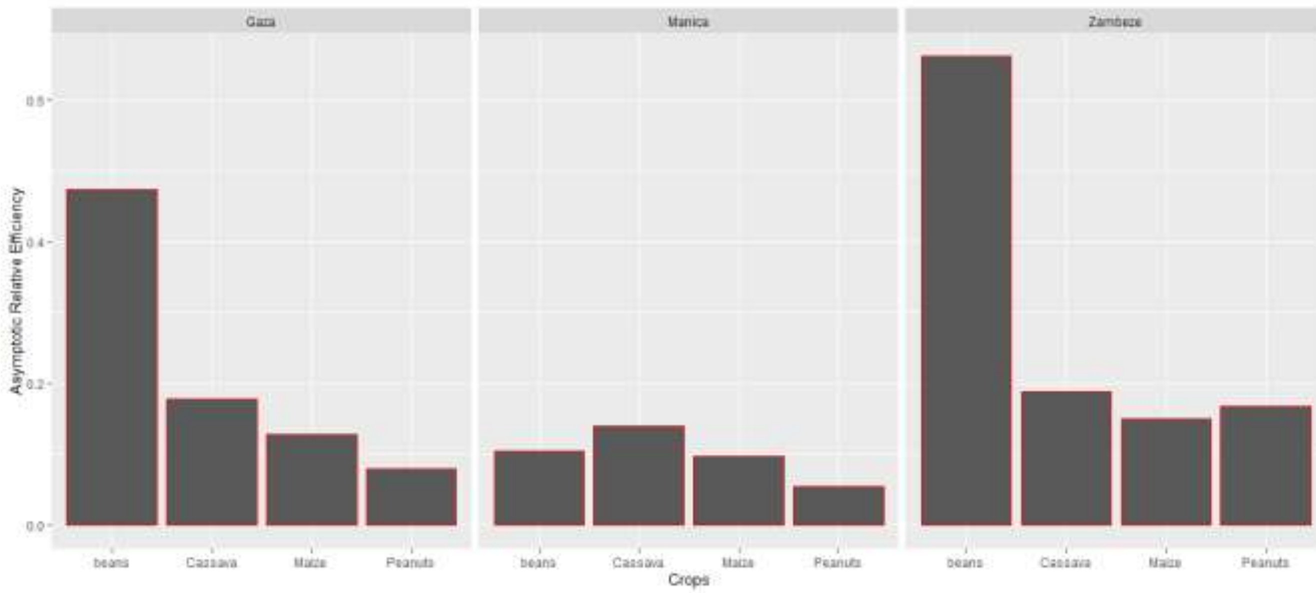

Figure 6. Asymptotic Relative Efficiency estimates of SFS approach and its counterpart IAI

The results of this study consistently demonstrate that the SFS approach generates substantial gains on the efficiency of crop cultivated area estimates when compared to its counterpart IAI. This is most clearly reflected in the values of Asymptotic Relative Efficiency (ARE), which decrease to approximately 5% under the SFS approach. This means that the variance of crop area estimates obtained using the SFS approach decreases significantly and represents up to 5% of the IAI variance estimates. The results of ARE of SFS compared to its counterpart IAI shows consistency that SFS approach generates better results. This trend is further corroborated by the analysis of other statistical performance metrics, including the range of confidence interval and the margin of error. In fact, when looking at the standard error estimates reported in Table 2, for all analyzed cases, it is clear that there is a substantial decrease in the standard error estimates for all crop areas estimates when the SFS approach is used. This is an indication of the increase in the efficiency of crop area estimates due to the decrease of the variance estimates.

These results highlight the strategic importance of leveraging the use of the SFS approach especially for crop area estimates. The cultivated area is a key variable to measure crop yields and using more reliable methodologies that measure crop area with more precision will contribute to improve agriculture production statistics. Furthermore, the SFS approach allows to overcome the challenge of measuring crop areas especially for intercropping systems which mostly characterizes the Mozambique's agriculture. In fact, the use of drones (UAV) during the data collection process to capture spatial images of cultivated plots, provides images that can easily be processed in a timely manner to estimate covered land per crop with more accuracy even in the intercropping systems. We indeed believe that crop cultivated area estimates are the key point of complementarity and synergies between the IAI and SFS approaches. While IAI is the official well-established methodology for agricultural data collection in Mozambique and it continues being an important source of agricultural information for government planning and decision makers, IAI methodology can also incorporate the SFS approach especially for crop cultivated area estimates purpose.

4. Conclusion

The findings from this study demonstrate developing sampling frames through spatial stratification in which the strata derivation takes into consideration different factors such as soil and elevation characteristics, land forms, weather conditions, level of farming intensity and NDVI profiles, offers a robust alternative methodology to estimate crop cultivated areas with more precision. This approach is particularly well-suited for intercropping production systems which are widely practiced by most of the smallholder farmers in Mozambique.

The application of the SFS methodology to estimate crop cultivated areas has shown improved efficiency when compared to the traditional methodology currently used in the Integrated Agricultural Survey (IAI).

The SFS approach presents a valuable complementary tool that can be integrated into the planning and implementation of future IAIs, enhancing both methodological rigor and operational effectiveness in agricultural data collection in Mozambique.

Acknowledgement

To the South Korean Government for the funding provided to conduct the study.

References

Casley, D.J., Kumar, K. (1998). The collection analysis, and use of Monitoring and Evaluation Data. Ram C. Malhotra, editor of the series. The Johns Hopkins University Press Baltimore and London. 188p.

CENACARTA (Mozambique National Cartography and Remote Sensing Centre). Landforms in Mozambique - 2012. www.cenacarta.com

Damodar N. Gujarati. 2009 Basic Econometrics (5th Edition). McGraw Hill.

Jeffrey M. Wooldridge. 2009. Introductory Econometrics: A modern approach (6th Edition). South Western.

MADER (2024). Inquérito Agrário Integrado - 2023. Maputo – Moçambique

Mood A.M., Graybill, F.A.,BOES, D.C. Introduction to the Theory of Statistics, 3rd Edition 1974.

Murphy, J., Casley, D.J., and Curry, J.J. (1991) Farmer's Estimation as a Source of Production Data. Methodological guidelines for cereals in Africa. World bank technical paper number 132.

Wackerly, D.D., W. Mendel, e R.L. Scheaffer. Mathematical Statistics with Applications, 6 th Edition, 2002.

Wang, Y.G.and Lin, X. (2005). Effects of Variance-Function Misspecification in Analysis of Longitudinal Data. Biometrics 61, 413-421.