Land area measurement bias: Evidence from West African countries

Anatole Goundan*
International Food Policy Research Institute, West and Central Africa Office, Titre 3396, Dakar Almadies, Senegal. E-mail: a.goundan@cgiar.org

Lea Vicky Magne Domgho
International Food Policy Research Institute, West and Central Africa Office, Titre 3396, Dakar Almadies, Senegal. E-mail: l.v.magnedomgho@cgiar.org

*Corresponding author

Abstract

Land area under farming is a cornerstone in agricultural statistics and policies. Crop yield or input use (fertiliser, pesticide, etc.) intensity is commonly used to compare countries over the world. Such comparison will not be relevant if land area under use is not well measured for one or several countries. In developing countries, farming lands are estimated based on the plot size reported by farmers during surveys. However, farmers' self-report is always subject to some bias due to lack of knowledge or strategic reasons. Therefore, it is essential to assess the direction of such bias to balance international comparison based on these statistics. Consequently, this paper uses data from the World Bank's Livings Standards Measurement Study – Integrated Surveys of Agriculture (LSMS-ISA) for four West African countries to quantify the nature and extent of land area measurement error. It also explores the sources of the bias in land measurement in these countries. Our findings indicate that, at country level, the relative bias (bias divided by the GPS measure) was estimated at minus 14% in Niger, minus 39% in Mali, 19% in Burkina Faso, and 87% in Nigeria. By plot size, the results show that, in general, farmers overreport for small plots and underestimate for larger plots. There is a systematic difference between self-reported and GPS measures in all countries. Therefore, the use of a more systematic approach and relatively accessible technology like GPS would improve agricultural statistics and allow a better international comparison.

Key words: land measurement; GPS measure; self-report; acreage gap; West Africa

1. Introduction

Agriculture is an important sector in developing countries, and billions of people depend on it exclusively for their income and subsistence. Therefore, agriculture is a strategic area of interest for policymakers in terms of food security, poverty alleviation, biosecurity, etc. The usual components of agricultural policies are based on agricultural land and productivity. As stated by Carletto *et al*. (2015:2), "Land is a key measure of absolute and relative farmer wealth, a critical input in production, and a key variable for normalizing agricultural input use and output measures." In addition, in order to compute crop production, agricultural statisticians use harvested area and yield (Keita & Carfagna 2009). Therefore, a good estimation of plots is essential and may have huge consequences for agricultural inputs and outputs.

During surveys there are different options for obtaining the land area that farmers are operating on. The easiest and 'costless' manner is to ask farmers to report directly or, should we say, to 'guess' the size of their lands. As said, it is the easiest way to estimate land under use quickly without adding additional time to the data collection process. It comes without 'costs' if farmers have quite good knowledge of land measurement. In developing countries, there is a priori no guaranty that farmers will have a good estimation of their plot sizes. In addition, there is a great deal of evidence that shows that the self-reported (SR) land area is usually biased (DeGroote & Traoré 2005; Carletto *et al*. 2013; Holden & Fisher 2013; Carletto *et al*. 2015). As an alternative to self-reporting, two other methods, namely Compass and Rope (CR) and the Global Position System (GPS), are used to obtain accurate and precise land area measurement. The first one is commonly considered to be the 'gold standard' measure (FAO 1982) and relies only on basic geometry, while the second one is based on advanced technology and offers a practical approach to the measurement of land area (Kelly & Donovan 2008). However, GPS and CR measures require careful implementation; indeed, GPS measurements may also be affected by the GPS device's environmental factors, and this can give rise to concerns of accuracy of measurement of very small plots. Although the CR method is highly accurate when performed properly, it is more cumbersome, costly and time consuming than self-reporting and GPS. Moreover, recent evidence shows that the gap between the GPS and CR measure is very small and that GPS could be used as a reliable alternative to CR (Carletto *et al*. 2013).

With new advanced data collection technologies, surveys use more and more GPS devices to report land size systematically. However, some field work continues to rely solely on farmers' self-report. Therefore, it is crucial to know how reliable the land area size is collected through such surveys. This paper is an attempt to assess the accuracy of land area self-reported size in different contexts, especially in West African countries. It analyses the distribution of the gap between SR and GPS measures, which may help us construct a typology of land measurement error in West African countries. It will also contribute to the identification of factors that may affect misreporting by farmers in different contexts.

The rest of this paper is organised as follows. The next section provides a description of the data and, especially, the discrepancies observed between self-reported area size and GPS measures. Section 3 succinctly presents the econometric models used and discusses the findings. The last section summarises the main findings of the analysis and some policy implications.

2. How widely do farmers misreport land area?

2.1 Data and GPS measure coverage

To analyse the existence of any gap that may be present between farmers' estimates of plot size and the measurements given by GPS devices, this study uses data collected under the Living Standards Measurement Study – Integrated Surveys of Agriculture (LSMS-ISA) project. This project covers four countries in West Africa, namely Burkina Faso, Mali, Niger and Nigeria. Each survey is representative at country level and has a special focus on agriculture. The data collection methodology is the same for all the countries. Regarding land measurement, farmers were asked to report their plot size, whereas enumerators also, where possible, utilised GPS devices to estimate the plot size.[1] Data were not collected for the same year everywhere. For example, in Burkina Faso, Mali and Niger, the data collection process took place from the post-planting period in 2014 to the post-harvest period in 2015, and in Nigeria it took place between 2012 and 2013. For this analysis, the data are limited to plots for which both measures were available. With that constraint, the total number of plots varies from 3 061 plots in Burkina Faso to 8 215 plots in Mali.

The coverage of GPS-measured plots varies widely from one country to another. In Burkina Faso, for example, only 11.9% of total plots have been measured by GPS, while in Niger, Mali and Nigeria we have complete data reaching 74.4%, 94.7% and 99.8% of plots respectively. This very low data

[1] With well-trained enumerators, GPS estimates provide an effective approximation of CR measurements at a lower cost.

completion in Burkina Faso is due to the fact that GPS devices were not available (about 99%).[2] Only 0.2% of the enumerators reported that plots were too small to be measured by GPS. Even at the regional level in Burkina Faso, the GPS coverage is very low (see Figure 1). Therefore, the results for Burkina Faso may not display the real situation of land measurement error in this country. Nevertheless, we can get some insights at sample level, which may be useful for policy makers to some degree.

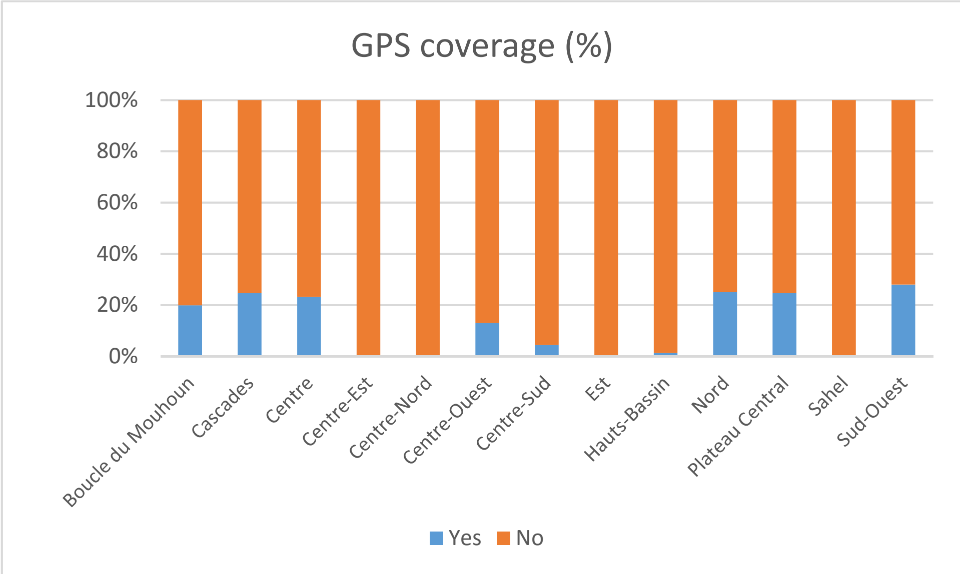

Figure 1: GPS coverage by region
Source: Author's calculation based on LSMS data

2.2 Difference between SR and GPS measures

Table 1 reports some descriptive statistics on land measurement issues. Two panels are considered, with the first one depicting information at plot level and the second one displayed at household level (aggregate plot information at household level). The rationale behind this is that households may misreport plot size at individual plot levels but have more accurate measurements at the aggregate level. Another level of analysis here is the plot size level. The reporting bias may depend on whether the plot is small or large. Therefore, data are split into five categories depending on the plot size. We follow the same categories as in Carletto *et al*. (2015). With these categories we can analyse measurement error for very small plots and for larger plots. The summary statistics are computed and presented for each country and for all countries together (pooled). To limit the influence of outliers, we excluded from our analysis all plots with a GPS (and SR) size of less than 0.01 acres on the one hand, and on the other hand we trimmed off the top 1% of SR and GPS measurements; this dropped about 2% of the overall sample.

The plot size distribution across countries shows that most of the farming takes place on small plots of land less than one acre in size. The share of plots with a size less than one acre varies from 17.5% in Niger to 66.3% in Nigeria. The share of plots with a size greater than five acres appears to be relatively small. It is about 3.4% in Nigeria, 10.7% in Burkina Faso, 20.5% in Mali and 26% in Niger. On the aggregate level, 57% of the plots are small, while only 8% are larger than five acres.

As shown in Table 1, on average the two methods produce distinct estimates of land area (t-test shows that SR and GPS are not identical), except in Niger, where the aggregate bias is not significantly different from zero. Carletto *et al*. (2015), using 2011 survey data for Niger, found a positive bias for

[2] The questionnaire requires enumerators to report the reason why a GPS measure was not filled in for a selected plot.

this country. This suggests that the misreporting scheme may not be constant over time, even in the same country or context. In the overall sample, the average size of plots using GPS is 1.96 acres, which is 0.62 acres larger than the area size reported by the farmers. Moreover, 69.2% of plots displayed a positive bias. In Burkina Faso and Nigeria, farmers on average declared plot sizes greater than the GPS measure by 0.41 acres and 1.04 acres respectively, while in Mali, farmers underreported their plot sizes by 1.70 acres. The rate of overreporting was 65.8% in Burkina Faso, 50.7% in Mali, 62.2% in Niger and 72.9% in Nigeria. Disaggregating by plot size (based on GPS measurements) revealed, as in Carletto *et al*. (2013) and De Groote and Traoré (2005), that there is a systematic difference between SR and GPS, even in Niger.[3] In addition, our data point out that the bias is much more pronounced in the tails (smaller plots are overreported, while larger plots are underreported for all countries), thereby showing that most of the bias lies there, which confirms the observation made by Carletto *et al*. (2015) on the data from Malawi, Uganda, Tanzania and Niger.

Regarding the relative bias, the results showed that the sign, as well as the magnitude of the bias, both seemed to be both to plot size. In general, small farmers – those cultivating less than 0.5 acres – systematically overestimate the size of their plots and exhibit much higher bias, and the degree of that overestimation declines as plot size increases. Carletto *et al*. (2013; 2015) and De Groote and Traoré (2005) observed the same trend in Malawi, Niger, Tanzania, Uganda and Southern Mali. When pool data are used, the relative bias stands at 382% of the average GPS measure for the lowest group, while for the highest group it is about minus 44.1%. A similar pattern is observed for individual countries. At the household level, the average gap is estimated at 1.2 acres in Burkina Faso, minus 7.3 acres in Mali, minus 0.00 acre in Niger, and at 1.9 acres in Nigeria. In general, the results found at plot level are confirmed at the household level, except in Burkina Faso, where all classes are overestimated.

3. Determinants of acreage discrepancy in West African countries

3.1 The econometric approach

The following section explores the factors affecting the acreage bias between GPS and self-reported measures. To do so, this study uses a simple OLS regression model. The model is presented as follows:

$$Y_i = \alpha + \theta_1 GPS_i + \theta_2 GPS_i^{\,2} + \beta \mathrm{X_i} + \varepsilon_i$$

where Y_i is one of our four alternative dependent variables, defined as follows:

- **Bias**: This is the simple difference between the SR measure and the GPS measure, in acres (SR - GPS). This specification helps us to isolate factors that increase (or decrease) the bias. In other words, it allows us to determine the direction of the effect of a selected variable on the bias.
- **Absolute value of bias**: The use of the absolute value of measurement error allows us to know how an explanatory variable influences the magnitude of the bias.
- **Bias when GPS < one acre**: The descriptive statistics in the previous section revealed that SR is systematically overreported for small plots, therefore our third specification seeks to uncover the drivers of bias observed for small plots. We consider as small plots all plots smaller than one acre in size.
- **Bias when GPS > five acres**: This specification helps to reveal what explains the reporting bias observed for large plots. We assume in this study that a large plot must have a plot size greater than five acres.

[3] A t-test was used to test the differences between the two measures. The results of these tests are available upon request.

Table 1: Land measurement gap by plot size

Country	Farm size	No. of plots	Plot mean (acres) GPS area	SR area	Mean bias (SR-GPS)	Bias as % of GPS area	Farm mean (acres) No. of farms	GPS area	SR area	Mean bias (SR-GPS)	Bias as % of GPS area
Burkina Faso											
	< 0.5	671	0.28	1.15	0.87	316.42	51	0.32	1.49	1.17	370.30
	0.5–0.99	513	0.73	1.51	0.78	106.54	60	0.76	1.74	0.98	128.78
	1–1.99	699	1.46	2.09	0.63	43.42	128	1.52	2.67	1.15	75.84
	2–4.99	843	3.14	3.58	0.44	13.87	291	3.42	5.13	1.71	50.08
	> 5	328	8.75	7.12	-1.63	-18.66	510	11.41	12.37	0.96	8.42
Total	**0.025–38.68**	**3 054**	**2.34**	**2.75**	**0.41**	**17.51**	**1 040**	**6.90**	**8.11**	**1.21**	**17.51**
Mali											
	< 0.5	1 301	0.25	0.93	0.68	266.77	63	0.23	2.25	2.02	870.28
	0.5–0.99	1 218	0.72	1.12	0.40	54.77	82	0.72	1.31	0.59	81.27
	1–1.99	1 544	1.46	1.59	0.13	9.18	105	1.42	3.68	2.26	159.58
	2–4.99	2 097	3.17	3.24	0.08	2.45	299	3.38	4.32	0.94	27.91
	> 5	1 801	16.02	6.61	-9.41	-58.74	1 389	26.53	15.79	-10.73	-40.46
Total	**0.012–89.8**	**7 961**	**4.57**	**2.87**	**-1.70**	**-37.26**	**1 938**	**19.52**	**12.25**	**-7.27**	**-37.26**
Niger											
	< 0.5	399	0.27	2.12	1.85	690.99	47	0.25	1.83	1.59	648.40
	0.5–0.99	367	0.76	2.81	2.05	270.15	53	0.75	3.20	2.45	325.28
	1–1.99	641	1.48	3.06	1.59	107.30	139	1.49	4.16	2.66	178.22
	2–4.99	1 314	3.19	4.67	1.48	46.39	481	3.44	6.84	3.40	98.93
	> 5	1 170	12.75	8.13	-4.62	-36.20	1 124	16.10	14.02	-2.07	-12.87
Total	**0.012–88.96**	**3 891**	**4.88**	**4.88**	**0.00**	**-0.02**	**1 844**	**10.86**	**10.86**	**0.00**	**-0.02**
Nigeria											
	< 0.5	2 211	0.22	1.05	0.83	380.97	702	0.25	1.33	1.09	442.02
	0.5–0.99	1 059	0.72	2.28	1.56	214.68	430	0.74	2.85	2.12	287.18
	1–1.99	1 068	1.41	3.01	1.60	113.25	593	1.44	4.19	2.74	189.85
	2–4.99	653	2.94	4.22	1.28	43.43	674	3.04	6.00	2.96	97.48
	> 5	190	8.81	5.53	-3.28	-37.22	318	9.31	9.15	-0.15	-1.66
Total	**0.01–38.96**	**5 181**	**1.15**	**2.18**	**1.04**	**90.29**	**2 717**	**2.15**	**4.10**	**1.94**	**90.29**
Global											
	< 0.5	4 582	0.22	1.07	0.85	381.96	863	0.25	1.35	1.10	446.08
	0.5–0.99	3 157	0.73	2.20	1.47	202.39	625	0.74	2.83	2.09	283.99
	1–1.99	3 952	1.42	2.84	1.42	99.63	965	1.45	4.17	2.72	188.10
	2–4.99	4 907	3.05	4.15	1.10	35.93	1 745	3.11	6.05	2.94	94.41
	> 5	3 489	12.24	6.85	-5.39	-44.07	3 341	14.91	12.16	-2.75	-18.46
Total	**0.01–89.8**	**20 087**	**3.68**	**3.13**	**-0.56**	**-15.11**	**7 539**	**9.81**	**8.33**	**-1.48**	**-15.11**

Note: 1) National data use sample weights, while the metadata (global level) do not; 2) Our analysis includes plots with both GPS and SR areas at least 0.01 acres (top 1% of GPS and SR areas trimmed); 3) CV here stands for coefficient of variation, computed as the standard deviation divided by the average bias; 4) Area of land at household level is just the aggregation of individual plots per household; 5) Global level here is the pooled data for all countries. Source: Author's calculation based on LSMS data.

The last two specifications make it possible to qualify the results from the first specification. GPS_i represents the GPS measure of plot i. As stated by Carletto *et al*. (2015), the inclusion of the GPS measure and its squared control for the influence of the 'real' plot size. If the coefficients associated are significantly different from zero, the bias is not independent of the plot size. If not, the observed bias is independent of the true area measurement (assumed here to be the GPS measurement). We consider additional control variables, X, which include plot-specific variables, and some household characteristics. The theory is that plot characteristics may affect both the accuracy of GPS and SR, while household-specific variables will only influence the SR measure. The plot-level variables are the plot's relief (plain, plateau and other, the reference being plain), the soil type (sandy, clay and other, where sandy soil is set as reference), the mode of acquisition of the plot by the household (heritage or not), and the distance from the plot to the household dwelling. In Nigeria, the relief and the soil type as defined previously are not available, but instead there is information on the plot's slope and elevation; that information was used for the country. At household level, we considered the gender of the household head (1 if female and 0 otherwise), his or her education (no level, primary school level, and at least secondary school level), age and its squared, and the total number of plots used by the household. If one or many control variables are significantly different from zero, it suggests that these variables can be used to predict the gap between the SR and GPS measures. We also controlled for spatial difference in the reporting by including regional fixed effects.[4]

3.2 Empirical results

This section presents results from the regression analysis, which aimed to identify the determinants of measurement bias, defined here as the difference between SR and GPS measures; the GPS measure was used here as the benchmark. For each of our four countries (Burkina Faso, Mali, Niger and Nigeria), we considered the four variants of the dependent variable as mentioned earlier, and the results are presented by model specification for all countries (see Tables 3 to 6). Robust standard errors have been used in all estimations to control for heteroskedasticity.

Table 2 reports the results for the first specification. The results show a very big explanatory power for Mali (89%) and Niger (73%), and a relatively small explanatory power for Nigeria (17%) and Burkina Faso (24%). A large explanatory power suggests that factors have a strong correlation with the dependent variable. For all countries under consideration, the plot size showed a nonlinear and negative correlation with the reporting gap. In Burkina Faso and Nigeria, this finding suggests that the bias is positive (overreporting) for small and moderate plot sizes, and negative for larger plots, since the constant term is significantly positive. For the latter countries, the threshold at which bias switched from positive to negative is 4.8 acres in Burkina Faso and 6.3 acres in Nigeria. For the other countries (Mali and Niger), the bias is systematically negative (underreporting) and small in absolute value for small plots, and very wide in absolute value for larger plots. Regarding other plot-specific variables, the results revealed that they do not all matter in all contexts. Plot topography (relief, soil type, plot elevation and slope) showed a significant effect in Mali, Niger and Nigeria. The mode of acquisition of plots (dummy variable being 1 if heritage) is associated with overreporting of plot size in all countries except Niger. This finding suggests that farmers who acquired their plot through inheritance tend to overestimate their plot size. Similarly, distance between plot and household dwelling was found to have a positive and significant effect on the bias in Niger. Therefore, households that are very far from their plots are found to display greater misreporting.

Concerning household-level variables, gender showed a significant and negative effect on the bias in Burkina Faso and Niger, suggesting that female-led households were more accurate in their reporting than male-led households. The age of the household head displayed a U-shaped relationship with the acreage bias in Mali and Niger. Thus, households led by younger persons tend to report with a positive bias, while other households were likely to be associated with negative bias. Education level also has a significant and negative impact on the bias in Burkina Faso, Niger and Nigeria. That means that the land measurement error is lower for more educated households. In Mali, education does not matter for the observed bias. Similar results were also found in Uganda by Carletto *et al*. (2013). This absence of impact may be the result of the fact that even literate people have little information on

[4] The models with and without regional fixed effects were compared and the best one was selected. The model with regional fixed effects was selected if at least one region had a significant coefficient. For each country, the following regions were considered: Centre, Sud-Est, Cascades, Plateau Central, Nord, Hauts-Bassins, Boucle du Mouhoun, Centre-Sud and Centre-Ouest in Burkina Faso; Kayes, Gao, Tombouctou, Mopti, Segou, Sikasso and Koulikoro in Mali; Agadez, Diffa, Dosso, Maradi, Niamey, Tahoua, Tillaberi and Zinder in Niger; and South South, North Central, North East, South East, North West and South West in Nigeria.

their plots, or that farmers do not regularly report their acreage for their own use. So, they are forced to guess the plot size. Regarding the number of plots under farming by household, the results show a strong negative linkage with the reporting bias for all countries except Burkina Faso, where there is no effect. This negative effect of the total number of plots reveals that households with many plots are less precise than those with fewer plots. This suggests that, in those countries, respondent fatigue during the data collection process led to misreporting.

Table 2: Determinants of area size bias

	Burkina Faso	Mali	Niger	Nigeria
GPS	-0.266*** (-6.535)	-0.519*** (-20.619)	-0.469*** (-14.918)	-0.231*** (-3.429)
GPS squared	-0.008*** (-3.175)	-0.006*** (-14.656)	-0.006*** (-14.274)	-0.025*** (-4.971)
Plot-specific variables				
Clay soil (ref. sandy)	0.027 (0.260)	-0.170 (-1.360)	-0.814*** (-3.108)	
Other soil (ref. sandy)	-0.047 (-0.392)	-0.378*** (-3.187)	0.086 (0.336)	
Plateau relief (ref. plain)	0.149 (1.288)	0.183 (0.933)	0.412** (2.006)	
Other relief (ref. plain)	0.075 (0.529)	0.205* (1.910)	0.248 (0.949)	
Plot slope (%)				-0.005 (-0.415)
Plot elevation (m)				-0.001*** (-3.580)
Plot was acquired through heritage	0.189** (-0.392)	0.424*** (-3.187)	-0.074 (0.336)	0.213**
Distance from HH (km)		0.007 (0.595)	0.134** (1.966)	0.004 (1.234)
Household-level variables				
HH is female	-0.251** (-2.156)	-0.348 (-1.198)	-0.966*** (-4.021)	-0.131 (-1.074)
Age	0.004 (0.255)	0.094*** (5.178)	0.073** (2.246)	-0.012 (-0.740)
Age squared	-0.000 (-0.333)	-0.001*** (-4.263)	-0.001** (-1.970)	0.000 (0.224)
Primary school level (ref. no level)	0.078 (0.717)	0.039 (0.354)	-0.232 (-1.211)	-0.174* (-1.881)
Secondary school level or more (ref. no level)	-0.844** (-2.005)	0.421 (1.171)	-0.695** (-2.255)	0.781 (1.317)
Number of plots	0.009 (0.873)	-0.071*** (-5.845)	-0.149*** (-4.999)	-0.132*** (-4.130)
Constant	1.449*** (2.814)	-2.211*** (-3.741)	-0.723 (-0.927)	2.414*** (5.223)
Observations	3 052	6 329	3 890	5 181
R-squared adjusted	0.236	0.893	0.728	0.174
Regional fixed effect	YES	YES	YES	YES

Robust t-statistics in parentheses; *** p < 0.01, ** p < 0.05, * p < 0.1

Table 3 reports results for the second specification, where the focus was on how each factor affects the magnitude of the measurement error (the dependent variable here is the absolute value of the bias). When the absolute level of the bias is considered, the relationship with plot size becomes monotonically positive, with a small curvature (quadratic term is significant) everywhere. Therefore, larger plots are associated with larger bias. The magnitude of the measurement error increases with the plot size. Clay soil is found to increase the absolute bias relative to sandy soil in Mali, while plateau plots in Burkina Faso are associated with a larger bias than plains plots. The mode of plot acquisition is only significant in Burkina Faso and reveals that the acquisition of a plot through heritage is a factor that increases bias. At the 10% level of significance, distance between plots and household also appears to increase bias in Niger, while it has the opposite effect in Mali, at the 5% level. Gender and education are two strong bias-reduction factors in Niger. In other countries, these two variables were non-significant. The total number of plots displayed some very contrasting effects from one country to another. In Burkina Faso and Nigeria there was a negative and significant impact, while in Mali the impact was significantly positive.

Table 3: Determinants of the absolute value of the area size

	Burkina Faso	Mali	Niger	Nigeria
GPS	0.206*** (6.487)	0.472*** (19.626)	0.245*** (8.761)	0.276*** (4.902)
GPS squared	0.008*** (3.721)	0.007*** (16.253)	0.009*** (20.855)	0.015*** (3.214)
Plot-specific variables				
Clay soil (ref. sandy)	-0.006 (-0.074)	0.243** (2.058)	-0.308 (-1.402)	
Other soil (ref. sandy)	-0.081 (-0.816)	0.016 (0.137)	-0.241 (-1.101)	
Plateau relief (ref. plain)	0.115 (0.900)	-0.145 (-1.389)	0.097 (0.426)	
Other relief (ref. plain)	0.075 (0.529)	0.205* (1.910)	0.248 (0.949)	
Plot slope (%)				-0.009 (-0.794)
Plot elevation (m)				-0.001*** (-3.580)
Plot was acquired through heritage	0.143* (-0.816)	0.060 (0.137)	-0.196 (-1.101)	0.100
Distance from HH (km)		-0.025** (-2.252)	0.109* (1.817)	0.004 (1.069)
Household-level variables				
HH is female	-0.108 (-1.182)	-0.122 (-0.508)	-0.426** (-2.104)	-0.044 (-0.380)
Age	-0.008 (-0.544)	-0.007 (-0.427)	-0.036 (-1.265)	-0.012 (-0.834)
Age squared	0.000 (0.772)	0.000 (0.039)	0.000 (1.408)	0.000 (0.381)
Primary school level (ref. no level)	0.000 (0.002)	-0.145 (-1.386)	-0.147 (-0.909)	-0.129 (-1.537)
Secondary school level or more (ref. no level)	0.395 (1.093)	0.236 (0.750)	-0.654*** (-2.756)	0.613 (1.065)
Number of plots	-0.014* (-1.685)	0.021** (1.961)	-0.011 (-0.410)	-0.118*** (-3.974)
Constant	0.939** (2.145)	0.665 (1.229)	1.164* (1.729)	2.191*** (5.159)
Observations	3 052	6 329	3 890	5 181
R-squared adjusted	0.301	0.904	0.750	0.175
Regional fixed effect	YES	YES	YES	YES

Robust t-statistics in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Previous model specifications revealed how each factor is associated with the direction of the bias (positive equals overestimation and negative relates to underestimation) and/or the magnitude of the bias (positive means higher absolute value of bias, and negative means lower absolute value of bias). Those two sets of regressions are not adapted to answer questions such as why small plots are systematically overestimated or larger plots are underreported by farmers, as shown in descriptive statistics. The next paragraphs deal with these issues.

Table 4 reports the results for the regression of the determinants of farmers' reporting bias (*SR-GPS*), with only small plots being considered (GPS measure less than 1 acre). Before going to the results, it is important to note that the GPS measure is not very accurate for small plot sizes. Authors have found that, on average, GPS measurement devices underestimate the area size of plots smaller than 1.24 acres (Schoning *et al*. 2005; Keita & Carfagna 2009). Therefore, the overestimation observed for small plots in general could have been exacerbated by the imprecision of GPS devices for those plots. The first observation from the results in Table 4 compared with previous results (Table 2 and Table 3) is the very small explanatory power, especially for Burkina Faso (from 24 to 30% to 6%), Mali (from 89 to 90% to 4%), and Niger (from 73 to 75% to 12%). In Nigeria, the explanatory power of these three specifications stayed quite stable, between 11% and 18%. This means that the factors under consideration, even though relevant for measurement error in general, do not explain the small plot's bias in Burkina Faso, Mali and Niger. For two of these three countries (Burkina Faso and Mali), even plot size terms did not show a significant effect on the bias. This result suggests that the measurement error was independent of the associated plot size as measured by GPS. Since the GPS measure may not be very accurate for small plots, these results are difficult to support. This calls for a better strategy in measuring plot size using GPS to guarantee its accuracy for small plots. Moreover, about 32% to 39% of plots surveyed in those countries are less than one acre in size. Consequently, it is important to consider this issue by paying more attention to ensure good data quality for agricultural policies. Another explanation of the fact that plot size (GPS) is not significantly different from zero is the one given by Carletto *et al*. (2015). According to these authors, this result suggests that any difference between farmer estimates and GPS measures could be controlled for completely by household or plot-level variables. For other countries, plot size does not appear to be independent of the reporting gap. In Niger, the measurement error had a strong non-linear (inverted U-shaped) relationship with the plot size. Therefore, in the context of Niger, small plots are not associated with positive bias. Finally, in Nigeria, the results showed a linear and positive relationship between measurement error and plot size. Therefore this finding confirms that self-reporting is overestimated for small land areas. Some determinants of the bias of small plots include the soil and relief of the plot (Mali and Niger), the mode of acquisition (Burkina Faso and Mali), gender and education (Niger), age (Mali), and the number of plots everywhere.

Table 4: Determinants of the area size bias for smaller plots (< 1 acre)

	Burkina Faso	Mali	Niger	Nigeria
GPS	0.194 (0.361)	-0.757 (-1.241)	2.237 (1.173)	1.258** (2.336)
GPS squared	-0.502 (-1.033)	0.377 (0.555)	-2.781* (-1.723)	-0.810 (-1.245)
Plot-specific variables				
Clay soil (ref. sandy)	-0.003 (-0.022)	-0.001 (-0.005)	-0.997*** (-3.088)	
Other soil (ref. sandy)	0.091 (0.693)	-0.162* (-1.680)	0.058 (0.136)	
Plateau relief (ref. plain)	0.115 (1.233)	-0.326*** (-2.704)	0.941** (2.022)	
Other relief (ref. plain)	0.338 (1.436)	0.004 (0.026)	0.292 (0.695)	
Plot slope (%)				-0.027** (-2.381)
Plot elevation (m)				-0.001*** (-3.580)
Plot was acquired through heritage	0.342*** (0.693)	0.215** (-1.680)	0.160 (0.136)	0.139
Distance from HH (km)		-0.004 (-0.523)	0.058 (0.737)	0.003 (1.052)
Household-level variables				
HH is female	0.003 (0.033)	-0.058 (-0.143)	-0.884** (-2.302)	0.015 (0.159)
Age	0.003 (0.276)	0.046*** (3.025)	-0.063 (-1.001)	-0.007 (-0.440)
Age squared	0.000 (0.232)	-0.000*** (-2.616)	0.001 (1.127)	0.000 (0.202)
Primary school level (ref. no level)	0.151 (1.563)	0.094 (0.700)	-0.610** (-2.274)	-0.078 (-0.850)
Secondary school level or more (ref. no level)	-0.173 (-1.186)	0.016 (0.115)	-1.208*** (-3.112)	0.593 (0.772)
Number of plots	-0.025*** (-2.954)	-0.034*** (-2.625)	-0.161*** (-3.879)	-0.142*** (-4.886)
Constant	0.734** (2.142)	-0.789* (-1.883)	1.985 (1.464)	1.422*** (3.031)
Observations	1 184	1 941	766	3 270
R-squared adjusted	0.0624	0.0443	0.122	0.114
Regional fixed effect	YES	YES	YES	YES

Robust t-statistics in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 5 reports results of the determinants of the bias for larger plots (GPS greater than five acres). Unlike the results for small plots, the explanatory powers here are the largest among all the model specifications. This could mean that certain factors are more likely to explain the bias for larger plots specification than others. For all countries, a negative association was found between plot size and the bias, with a small curvature (a significant quadratic term for Mali and Niger). Whatever country is considered, the larger plots, on average, were correlated with underreporting. These results confirm those presented previously through the descriptive statistics. Apart from the plot size, some of the factors considered displayed a significant effect on the bias for larger plots. These include soil type and the number of plots (Mali), plot elevation and mode of acquisition (Nigeria), distance to plot and education (Burkina Faso), gender (Niger and Nigeria), and age (Mali and Niger).

Table 5: Determinants of the area size bias for larger plots (> 5 acres)

	Burkina Faso	**Mali**	**Niger**	**Nigeria**
GPS	-0.455** (-2.313)	-0.786*** (-19.948)	-0.617*** (-10.388)	-0.595** (-1.995)
GPS squared	-0.003 (-0.601)	-0.003*** (-5.757)	-0.005*** (-6.367)	-0.010 (-0.956)
Plot-specific variables				
Clay soil (ref. sandy)	0.386 (0.626)	-0.547 (-1.636)	0.174 (0.183)	
Other soil (ref. sandy)	0.091 (0.693)	-0.162* (-1.680)	0.058 (0.136)	
Plateau relief (ref. plain)	0.694 (0.923)	-1.160** (-2.409)	0.842 (1.333)	
Other relief (ref. plain)	0.338 (1.436)	0.004 (0.026)	0.292 (0.695)	
Plot slope (%)				-0.087 (-0.535)
Plot elevation (m)				-0.001*** (-3.580)
Plot was acquired through heritage	0.569 (0.923)	0.527 (-2.409)	0.382 (1.333)	1.676** (2.125)
Distance from HH (km) > 1 km	-3.140** (-2.457)			0.123 (0.948)
Household-level variables				
HH is female	0.860 (0.547)	1.517 (0.962)	-1.056* (-1.720)	-3.101* (-1.785)
Age	0.117 (1.236)	0.154*** (2.737)	0.123 (1.554)	0.199 (1.273)
Age squared	-0.001 (-1.479)	-0.001* (-1.899)	-0.001* (-1.672)	-0.002 (-1.329)
Primary school level (ref. no level)	0.320 (0.517)	0.410 (1.109)	-0.509 (-0.875)	-0.836 (-1.047)
Secondary school level or more (ref. no level)	-4.783*** (-3.317)	1.496 (0.884)	0.803 (0.914)	
Number of plots	0.177 (1.461)	-0.165*** (-4.291)	-0.119 (-1.088)	-0.261 (-0.662)
Constant	0.468 (0.156)	-1.481 (-0.905)	-0.971 (-0.365)	-0.443 (-0.094)
Observations	326	1 503	1 170	190
R-squared adjusted	0.302	0.924	0.802	0.426
Regional fixed effect	YES	YES	YES	NO

Note: Robust t-statistics in parentheses; *** p < 0.01, ** p < 0.05, * p < 0.1

4. Conclusion and recommendations

The objectives of this paper were to assess the importance of plot land area estimation bias by farmers in the context of countries in West Africa and to identify factors that influence acreage discrepancy between farmers' self-reported and GPS-measured plot size. The second objective was to highlight some elements that could help control the SR bias in future analyses.

The analysis for West African countries revealed a systematic discrepancy between SR and GPS land areas. On average, reporters overestimated land areas in Burkina Faso and Nigeria, whereas the plot size was underreported in Mali and Niger. The results also show a consistent impact of plot size on the quality of the estimation. As found by De Groote and Traoré (2005) and Carletto *et al.* (2013; 2015), small plot areas are generally overreported by farmers, while larger plot areas are systematically underreported. For example, in Burkina Faso, small-scale plots (less than one acre), which account for 39% of the plots, were overestimated by around 179%, while larger plots (more than five acres, 11% of the plots) were underestimated by 31%. In Mali, smaller plots (32%) were overreported by 106%, and larger plots (23%) were underestimated by 57.4%. In Niger, the results are similar, with small plots reporters (20%) overestimating by 312%, and large plots holders (30%) underestimating by 42.9%. In Nigeria, small-scale plots (63%) were largely overreported by 295%, while large-scale plots were underreported by 38.4%. Therefore, it is crucial for policymakers to think about this issue more seriously. These results suggest that the deviation of land areas bias is context specific, depending on circumstances within a country.

The sources of misreporting vary immensely from one country to another, and seem to be related to the 'real plot size', except for small plot sizes in Burkina Faso and Mali. Moreover, plot and household factors, as well as GPS, are more likely to explain bias for larger plots than others. Therefore, it appears difficult to control SR bias for smaller plots using these factors. It is worth noting that plot- and household-level variables affect land measurement error differently from one country to another.

Since there is a systematic gap between SR and GPS, it would be relevant to use the GPS to systematically measure land area size in future surveys. This would improve the quality of agricultural statistics, which are essential to better inform policy decisions. It would also be good to design policies or strategies to strengthen the capacity of farmers to better estimate the size of their activity. These strategies would help them hold their own operating accounts, assess the dynamics of productivity (crop yield), and finally report more accurate data (especially plot size), which again would contribute to better agricultural statistics. As some factors have a significant effect on the bias, future studies should use those factors (plot and household level) to control for SR bias when GPS measures are not available.

References

Carletto C, Gourlay S & Winters P, 2015. From guesstimates to GPStimates: Land area measurement and implications for agricultural analysis. Journal of African Economies 24(5): 593–628.

Carletto C, Savastano S & Zezza A, 2013. Fact or artifact: The impact of measurement errors on the farm size–productivity relationship. Journal of Development Economics 103: 254–61.

De Groote H & Traoré O, 2005. The cost of accuracy in crop area estimation. Agricultural Systems 84(1): 21–38.

FAO, 1982. Estimation of crop areas and yields in agricultural statistics. FAO Economic and Social Development Paper no. 22. Rome: FAO.

Holden ST & Fisher M, 2013. Can area measurement error explain the inverse farm size productivity relationship? Centre for Land Tenure Studies Working Paper 12/13, Norwegian University of Life Sciences, Ås, Norway.

Keita N & Carfagna E, 2009. Use of modern geo-positioning devices in agricultural censuses and surveys: Use of GPS for crop area measurement. Bulletin of the International Statistical Institute, the 57th Session, Proceedings, Special Topics Contributed Paper Meetings (STCPM22), 16–22 August, Durban.

Kelly V & Donovan C, 2008. Agricultural statistics in Sub-Saharan Africa: Differences in institutional arrangements and their impacts on agricultural statistics systems – A synthesis of four country case studies. Michigan State University International Development Working Paper No. 95, East Lansing, Michigan.

Schoning P, Apuuli JBM, Menyha E & Zake-Muwanga ESK, 2005. Handheld GPS equipment for agricultural statistics surveys: Experiments on area-measurement and geo-referencing of holdings done during fieldwork for the Uganda Pilot Census of Agriculture. Statistics Norway Report 2005/29.