A review of methods for estimating yield and production impacts

Andrew Dorward and Ephraim Chirwa
December 2010

Summary

This paper documents methodological lessons from experience in estimating yield and incremental production benefits from the Malawi Farm Input Subsidy Programme (FISP). Critical issues raised concern difficulties in obtaining reliable estimates of smallholder maize production, areas and yields with and without fertiliser. Comparison of methods and findings across a number of studies suggests that there is a significant upward bias (or overestimate) in area estimates obtained by relying on information from farmers on the areas of their crops. Use of GPS technology appears to provide an affordable but more reliable alternative method for measurement of plot areas, but further investigation is advisable regarding its accuracy and possible bias in measuring small plot areas.

Estimates of yield from information on total plot harvests are affected by bias in area measurement and, for given estimates of production per plot, over-estimates of area lead to under-estimates of yield. However these yield estimates are also dependent upon the accuracy of farmer estimates of plot production, and these are potentially prone to (a) errors in farmer estimates of production and harvest, (b) unwillingness of farmers to reveal total harvest, and (c) errors in reporting harvests in standard units. While (b) is likely to lead to under-estimates of harvest and yield, (a) and (c) might lead to over- or under- estimates of harvest. It is not clear what overall bias this might lead to. The main alternative method of yield estimation (enumerators harvesting yield sub plots) is costly and widely considered to lead to over-estimates of yield.

Estimates of crop yield response to fertiliser application are not affected by bias in plot area estimation if yield is estimated from reported whole plot harvest. However fertiliser responses are over- (under-) estimated if whole plot harvest is over-(under-) estimated by farmers. Fertiliser responses are also over-estimated if yield sub plots over-estimate yield, and this is exacerbated if plot areas are over-estimated, as this leads to an under-estimate of fertiliser application rate.

Over-estimates of fertiliser yield responses are also likely if analysis does not allow for the effects of early planting, plant density, seed type, number of weedings and early weeding: these management practices raise yields but tend to be correlated with fertiliser use. Analysis of subsidy impacts needs to separate out the effects of these practices and also estimate how far management practices are changed by receipt of subsidised inputs.

The following recommendations are highlighted regarding area and yield survey methods:

- IHS2 estimates of crop areas and yields should be treated with caution as they are likely to have over-estimated areas and under-estimated yields
- Further work is needed to investigate the accuracy of GPS area measurement for small plots, but GPS methods are preferable to farmer estimates of plot areas
- Continued work is needed to develop accurate methods for estimating plot yields
- Investigation of fertiliser yield responses may need to rely on formal trials to address problems of multi-collinearity across management practices
- Due attention should be paid to both inter cropped and pure stands in analysis and reporting of yield and area estimates

A review of methods for estimating yield and production impacts

Andrew Dorward and Ephraim Chirwa
November 2010

1 Introduction

This paper documents methodological lessons from experience in estimating yield and incremental production benefits from the Malawi Farm Input Subsidy Programme (FISP). The paper draws heavily on issues raised in Dorward and Chirwa (2010a) but also draws on new information from the National Census of Agriculture and Livestock (NACAL NSO, 2010) released later in 2010. There are significant difficulties in obtaining reliable estimates of smallholder maize production, areas and yields with and without fertiliser.[1] Methodological difficulties arise in obtaining accurate data on these variables and determining the yield impacts of changes in input use as a result of the programme.

Following this introduction the report summarises the main issues raised by various studies and compares their findings. It concludes with a summary of the challenges raised and possible ways forward for improving information on programme production impacts.

2 Lessons from various studies

In this section we describe the principle methodologies and results reported in a variety of studies of crop areas, production, and yield in Malawi in recent years. These are summarised in table 2.2 at the end of the section, which highlights a range of inconsistencies across different methodologies in different studies. We begin, however, by detailing in table 2.1 possible errors in yield estimation with different methods, as set out by Dorward and Chirwa (2010a).

Four different types of potential sources of estimation errors are considered in Table 2.1:

- random errors which are not likely to introduce bias,
- errors which may introduce bias in the results but the nature of the bias cannot be predicted,
- errors which are likely to introduce positive bias (overestimating yield and yield effects of different crop management practices), and
- errors which are likely to introduce negative bias (underestimating yield and yield effects of different crop management practices).

Possible means of reducing errors are identified for each potential source of estimation errors. These involve specific attention in survey design, implementation and analysis.

[1] Only maize is considered here as it has been the major focus of the input subsidy programme and its evaluation. Area, yield and production estimates for other crops face similar challenges as those considered here, with further challenges in harvest measurement for root crops and in allowing for intercropping.

Table 2.1 Yield estimation approaches and their errors and bias

Approach	**Farmer report on whole field harvest**		**Measurement of yield from 50m^2 yield sub plots**	
Methods	Yield is calculated from farmer reports of estimated harvest from each plot (using farmer defined units) divided by farmer estimates of the area of the plot (using farmer defined units). Applied to all plots cultivated by all sample farmers.		Yield is harvested & weighed from a 50m^2 yield sub plot (the ysp). The ysp is marked out by enumerators in the middle of the season for one maize plot for each of a subsample of farmers. Yield is harvested either by the farmer or by the enumerator & recorded by the enumerator.	
Possible errors	*Description*	*Possible remedial action*	*Description*	*Possible remedial action*
Principal sources of random errors	• Enumeration quality, farmer estimates of area & harvest • Small plots may have high % errors.	Survey & questionnaire design. Enumerator training & supervision Remove small plots from analysis	• Generally smaller sample size • Enumeration quality • Within field variability	Survey & questionnaire design. Enumerator training & supervision Gather more information specific to YSP management
Principal sources of errors with possible positive bias (overestimate yield &/or yield effects)	• Farmers may over report harvest to impress. Possible bias (overestimate) in harvested units • Correlation between variables (eg seed type & fertiliser) may bias estimates of their impacts	Enumerator training, supervision & interviewing Improve estimates of conversion coefficients for farmer units, estimate separately for different harvest units Selection of variables & estimation methods	• Enumerators may not site ysps randomly in parts of plot with low yield. • Farmers may include harvest from outside ysp • Fertiliser response may be overestimated with plot areas & fertiliser application rates over- and under-estimated respectively • Correlation between variables (eg seed type & fertiliser) may bias estimates of their impacts • High moisture at harvest/weighing	Enumerator training & supervision Estimate separately for enumerator & farmer harvest Use alternative plot area estimation methods – eg GPS Selection of variables & estimation methods Measure moisture content for sample
Principal sources of errors with possible negative bias (underestimate yield & yield effects)	• Farmers may under report harvest due to harvesting of green maize, unwillingness, to signal poverty, storing/ consuming &/or selling in small & non standard units, or very full bags. Possible bias (underestimate) in harvested units • Over estimate of plot areas	Enumerator training, supervision & interviewing Improve estimates of conversion coefficients for farmer units, estimate separately for different harvest units Use alternative plot area estimation methods – eg GPS (NACAL, IHS3)	• Farmers may harvest some yield before ysp harvest, without records.	Enumerators' pre- and post-harvest communication with farmers, observe & record evidence of harvesting (eg missing cobs)

Source: adapted from Dorward and Chirwa (2010a)

2.1 AISS2: the 2008/9 subsidy evaluation

Household surveys of coupon access and related agricultural activities were conducted in 2006/7 (AISS1) and 2008/9 (AISS2). These are generally considered to have provided reliable information on coupon access and use, but analysis of crop area, yields and production revealed serious difficulties with the data on these topics. In 2006/7 estimates of crop areas were broadly consistent with those for the previous IHS2, but yields were low and showed no increase. Analysis of seed and fertiliser impacts did not yield consistent results, and the data were not used. In the process anecdotal reports emerged regarding serious concerns in the analysis of the agricultural modules of the second integrated household survey (IHS2), although analysis had nevertheless continued and was released in the Poverty Vulnerability Analysis (PVA: Malawi Government and World Bank, 2006).

A number of changes were therefore made for the 2008/9 survey with a different field survey team with more stringent management, more attention to training and use of local measures for harvest, and the introduction of yield sub-plots on a sub sample of plots. Enumerators also carried out very approximate area estimates for the plots in which 50m^2 yield sub plots (YSPs) were laid, using a pacing method. These changes reflected a view that the major problems experienced in the 2006/7 survey were with the recording of plot production. However a large number of problems were again experienced during data analysis, with very low yields estimated from farmer harvest methods, but very high yields estimated from yield sub plots (see table 2.1). Both methods for measuring production and yield provided estimates of returns to fertiliser use, but these varied substantially with (a) the overall yield and (b) the variables included in regression models, indicating both substantial effects of a range of management practices (number and timing of weedings, time of planting, organic fertiliser use, time of fertiliser application and plant density as well as seed type and rainfall) and substantial multi-collinearity between the fertiliser application rates and other management practices, making it very difficult to reliably estimate the returns to fertiliser alone (Dorward and Chirwa, 2010a). In addition farmer area estimates of area were some 30% higher than enumerator estimates. This led to substantially increased yield estimates for these plots. It did not however affect estimates of returns to fertiliser use as it also increased estimated rates of fertiliser use. For the YSP estimates, however, yield was unaffected but returns to fertiliser fell (due to higher estimated fertiliser rates).

These issues were examined with the construction of a national maize balance sheet with different estimates and assumptions drawing information and comparing estimates from a range of different sources – an approach that is used again in this report, with new information, particularly from NACAL. It was concluded that no existing low cost survey methods are available for estimating fertiliser productivity on maize, and that further attention should be paid to this matter, as is being pursued in this report.

2.2 National Census of Agriculture and Livestock

The National Census of Agriculture and Livestock (NACAL) was conducted in the 2006/7 season, but the report was not available until the second part of 2010. It surveyed a nationally representative sample of 25,000 households with three rounds of visits per household (January, June and September 2007). Plot areas were measured using GPS methodology, not farmer estimates. Yields were estimated in three ways, by asking farmers for estimates of total production per plot (as in the IHS2, AISS1 and AISS2), for all main maize varieties a 7*7 metre sub plot was selected for harvesting, and for one household per enumeration area a maize plot was selected for full harvest by the enumerator. These methods are described in annex 2 of the NACAL report. This annex also contains a short review of studies on different methods. This complements a similar review in Dorward and Chirwa (2010a), reporting that yield sub plots tend to over-estimate yield by around 30%, and that farmers' estimates of production were generally more reliable provided that there are “reliable conversion factors which translate farmers’ traditional volumetric units into standard weight units”. However, it is subsequently stated that farmers reported maize yields in kilograms, and these are not traditional volumetric units. As regards estimates of plot areas, it also reported findings from studies across Africa that farmers seriously overestimate their plot areas. This does not affect estimates of total production but leads to under-estimates of yield per ha with famer reported area estimates, again in line with the findings reported for AISS2 by Dorward and Chirwa (2010a).

The main report presents estimates of crop areas and yield only for pure stand crops, but does present production estimates for maize as a whole. All production and yield figures presented are derived from farmer harvest estimates, not YSP estimates. The decision to use farmer harvest estimates is explained in annex 2 of the report where there is a comparison of pure stand yield estimates from the farmer estimates of plot production against YSP estimates, with YSP yield estimates on average 15% higher. However considerable variation is observed in the relation between farmer reported harvest estimates and YSP estimates, with a very low correlation coefficient. Analysis by seed type and ADD gives a number of situations where the YSPs provide lower yield estimates. The report concludes that the YSP method is less reliable because it is more complex for enumerators to implement. There is, however, no consideration of the difficulties in obtaining standardised conversion rates for maize when farmers report production in kg for maize that may have been harvested in stages (green and then dry) with the dry harvest commonly stored on the cob[2]. The NACAL report provides no information on the results from the small sub-sample of plots with full harvest by enumerators. As regards impacts of fertiliser application, weeding and seed type, pure stand yields are reported by number of weedings and number of fertiliser applications for each seed type, but there is no information on the effects of weeding/ fertiliser interactions, or on the yield effects of plant population, intercropping, or timing of planting or weeding.

2.3 Ministry of Agriculture and Food Security

The Ministry of Agriculture and Food Security (MOAFS) provides annual estimates of crop areas, yields and production for all the major crops in Malawi. Estimates are made by extension staff, based on a sample of plots, and information on these plots is aggregated up by district and Agricultural Development Division to provide national estimates. As compared with other studies, these tend to have higher estimates of the number of farm families, of maize yields, of the proportion of land under improved varieties, and of total production. Conversely, estimates of maize area per household and of land under local maize tend to be lower than in other studies and falling.

2.4 Chibwana et al.

Chibwana et al. (2010) analyse the effects of the subsidy programme on a sample of 380 farms from a three round (2002, 2006 and 2009) sample of farms from two districts. No information is provided on methods used in estimating area, production or yields. However estimates are presented of fertiliser effects derived from models that do not consider possible effects of multi-collinearity with other management variables.

[2] The NACAL annex also reports a high correlation between farmers’ pre-harvest estimates of plot production and post-harvest reports of production. This is taken as evidence of reliability of the pre-harvest estimates. It could however also be taken as evidence that the post harvest reports are, like the pre-harvest estimates, based on a general assessment of production rather than specific quantitative observation.

2.5 Holden and Lunduka

Holden and Lunduka (2010) report on yields and yield effects of the subsidy programme from surveys of 378 households carried out in six districts in 2006, 2007 and 2009. Plot areas were measured using GPS, while famers estimated plot production. Lunduka (pers comm.) has reported that GPS measurements give smaller plot areas than those reported by farmers. Estimates of average yields and maize areas per household are provided.

2.6 Comparison of studies

Table 2.2 sets out area, yield and national production estimates calculated with information from the different studies discussed above. Different reports provide different information (and have different gaps) reflecting different interests and methods used.

National crop areas are generally estimated from areas per household and number of households. There is some controversy over the number of farm households (see for example Dorward and Chirwa 2010b) with much larger figures estimated by the MoAFS as compared with the NSO. The IHS2 and AISS2 (farmer reported area) estimates of national area under maize are greater than those of the MoAFS despite smaller household numbers because there are larger estimates of area per household. However, this is not the case where the AISS2 area per household is adjusted downwards in line with enumerator estimates. The adjusted AISS2 estimates are then much closer to the NACAL estimates, which are lower than MoAFS estimates. Despite this, the MoAFS' larger estimated number of households leads to a smaller area per household. Holden and Lunduka estimate larger household maize areas, but their sample does not attempt to be nationally representative and may under represent areas with smaller holding sizes. We should expect the NACAL and AISS2 enumerator adjusted estimates to be the most reliable given (a) NACAL's use of the more objective GPS method and (b) it giving smaller area estimates, consistent with the tendency discussed earlier for farmers to overestimate areas.

We now consider yields and production. As discussed earlier (see table 2.1), yields estimated from farmer harvest reports are liable to be underestimates when plot areas are over-estimated by farmers. It is not clear what further bias may be caused by errors in farmers' reports of plot production. These may involve omissions from harvesting of green maize, farmers wishing to demonstrate the difficulties of farming, or farmers not wishing to report all their harvests. Farmers may also over- or under-estimate weights of maize when reporting kilograms harvested.[3] YSP estimates, on the other hand, are more likely to be over-estimates, though it is difficult to judge the extent of this. These expectations are broadly consistent with results from the different studies. IHS2 and AISS1 yields based on farmer reported areas are low compared to those estimated by other studies. The AISS2 yields with enumerator adjusted area estimates are still lower than NACAL and MoAFS yield estimates, which are comparable for local maize (and lower than Holden and Lunduka's estimates). However, NACAL's and Holden and Lunduka's estimates for hybrid maize are considerably lower than MoAFS estimates. Differences between NACAL farmer harvested and YSP estimates are, on average, not that large (though as commented earlier they are inconsistent when considered at a lower level of aggregation). The difference for the AISS2 YSP results is much larger (although some of these differences may also be caused by sampling as the YSPs plot samples were more likely to be on better managed plots).

[3] It is possible that farmers who tend to sell maize under-estimate the number of kilograms they produce while farmers who tend to buy maize over-estimate the number of kilograms they produce, if traders do not use standard measures when buying and selling grain. In addition, farmers that neither sell nor buy maize may either understate or overstate due to lack of experience with measurements.

Table 2.2: Comparison of summer maize production, area and yield estimates from different studies

	IHS2	AISS2			NACAL		MoAFS			Holden & Lunduka				Chibwana et al	
	2003/4	2008/9	2008/9	2008/9	2006/7	2006/7	2003/4	2006/7	2008/9	2005/6	2006/7	2008/9	All	2002/6/9	
Area	Farmer	Farmer	Enumerator	YSP	GPS	YSP	MoAFS	MoAFS	MoAFS	GPS	GPS	GPS	GPS	na	
Production	Farmer	Farmer	Farmer	YSP	Farmer	YSP	MoAFS	MoAFS	MoAFS	Farmer	Farmer	Farmer	Farmer	na	
Farm households ('000)	2,326		2,500		2,666			3,282	3,673						
National Areas (ha)					all maize										
local	848,735	1,135,000	873,077		932,932		736,901	575,014	507,294						
OPV	96,111	297,500	228,846		133,030		340,760	599,793	495,827						
recycled					169,553										
hybrid	687,772	830,000	638,462		421,867		459,990	511,635	439,056						
total maize	1,632,618	1,888,500	1,452,692		1,550,884		1,537,651	1,686,442	1,442,177						
Area per hh (ha)					all maize										
local	0.36	0.45	0.35		0.35			0.18	0.15						
OPV	0.04	0.12	0.09		0.05			0.18	0.15						
recycled					0.06										
hybrid	0.30	0.33	0.26		0.16			0.16	0.13						
total maize	0.70	0.76	0.58		0.58			0.51	0.44				0.71		
Maize % cult area	68%	74%			69%					73%	67%	64%	68%		
National Production (MT)															
local	421,810	612,900	612,900	1,349,777	865,701	908,986	508,809	534,321	524,486						
OPV	51,620	Na	na	na	203,614	244,337	399,255	1,190,571	1,108,559						
recycled					168,659	202,391									
hybrid	418,811	713,800	713,800	1,607,646	760,468	912,562	825,060	1,493,951	1,452,877						
total maize summer crop	892,241	1,269,072	1,269,072	2,903,932	1,998,442	2,268,275	1,733,125	3,218,843	3,085,922						
Yield (kg/ha)														subsidy	no
local	497	540	702	1,546	928	974	690	929	1,034	1,117	1,582	1,681	1,451	1,312	1,063
OPV	537	na	Na	na	1,531	1,837	1,172	1,985	2,236						
recycled					995	1,194									
hybrid	609	860	1,118	2,518	1,803	2,163	1,794	2,920	3,309	1,442	1,846	2,045	1,774	1,510	1,389
total maize	547	672	874	1,999	1,365	1,570	1,127	1,909	2,140	1,281	1,700	1,867	1608.71		
ratio hybrid/local	123%	159%	159%	163%	194%	222%	260%	314%	320%	129%	117%	122%	122%	115%	131%

Sources: as in main text. NACAL yield and production estimates are approximate calculations from pure stand and YSP differentials.
Areas may not sum to the total due to intercropping across maize varieties, but intercropping across maize varieties is relatively unimportant.

A final comment is needed as regards inconsistent treatment between studies as regards intercropping. Some studies do not appear to report area and yield separately for pure stand and intercropped plots, while NACAL only provides area and yield information on pure stand plots, with production (bit not area or yield information) on all plots (pure stand and inter-cropped).

We conclude from our consideration of the problems faced by and results from these different studies that GPS area measurements probably provide reasonably reliable information on plot and crop areas. Yield estimation is, however, much more problematic due to unresolved difficulties with both farmer reports of harvests and YSP harvests. Estimation of the yield impacts of fertiliser are further complicated by problems of multi-collinearity – and this is a problem that appears to receive almost no attention in most analyses. Ratios of hybrid maize yields to local maize yields also vary widely between different studies, ranging from around 1.2 (IHS2, Holden and Lunduka, Chibwana et al.) to around 3 (MoAFS), with AISS2 and NACAL giving intermediate estimates of around 1.6 and 2 respectively.

3 National maize areas, yields and production / consumption budgets

We conclude our consideration of maize areas and yields with a reconsideration in table 3.1 of national maize area, yield and production / consumption budgets presented in Dorward and Chirwa (2010a). These budgets describe alternative scenarios for low, medium and high maize production (and consumption) with for each scenario three different population estimates (a 'low' estimate from 2008 census figures, a 'medium' estimate, and a 'high' estimate using 2008/9 MoAFS figures). A budget is built up for total supply and consumption for each scenario. Supply is made up of smallholder and estate production and net imports while consumption allows for low calorie consumption per person and high importance of maize in the diet (in accordance with the low nutritional status of many in Malawi and a high proportion of cultivated area under maize) and higher consumption under conditions of higher production and lower prices. The first line in the table shows (***in bold italics***) smallholder maize yields that balance national supply with demand, that is yields that are consistent with all the other assumptions in the table.

This table differs from that presented in Dorward and Chirwa (2010a) as it uses the AISS2 enumerator adjusted estimate of maize area per household (0.58ha/hh) instead of a higher figure of 0.89ha/hh). This leads to higher estimated yields per ha to balance supply and demand – a rise from around 1,000 to 1,700 kg/ha in low and high yielding years, respectively. These higher yields are broadly compatible with the NACAL YSP and Holden and Lunduka yield estimates (see table 2.1). They are, however, lower than the MoAFS and AISS2 YSP yield estimates and higher than the IHS2, farmer reported AISS2, and (to a lesser extent) Chibwana *et al.* yield estimates.

Table 3.1 National Maize production and consumption budgets, 2008/9

Production Scenario	Low yield			Medium Yield			High yield		
Farm hh (million)	Low (NSO) 2.50	Medium 3.09	High (MoAFS) 3.67	Low (NSO) 2.50	Medium 3.09	High (MoAFS) 3.67	Low (NSO) 2.50	Medium 3.09	High (MoAFS) 3.67
PRODUCTION									
Yield (kg/ha)	***1,040***	***1,005***	***980***	***1,400***	***1,320***	***1,270***	***1,865***	***1,725***	***1,630***
area maize/household (ha)	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58
Total area (million ha)	1.45	1.79	2.13	1.45	1.79	2.13	1.45	1.79	2.13
Production/hh (kg)	603	583	568	812	766	737	1,082	1,001	945
Total Smallholder Production (million MT)	1.51	1.80	2.09	2.03	2.36	2.71	2.70	3.09	3.47
Estate production (million MT)	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
Total production (million MT)	1.66	1.95	2.24	2.18	2.51	2.86	2.85	3.24	3.62
TRADE									
Informal imports (million MT)	0.15	0.15	0.15	0.10	0.10	0.10	0.04	0.04	0.04
Formal imports /out of store (million MT)	0.13	0.13	0.13	0.00	0.00	0.00	0.00	0.00	0.00
Formal exports/ into store (million MT)	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.20	0.20
Net imports / out of store (million MT)	0.28	0.28	0.28	0.10	0.10	0.10	-0.16	-0.16	-0.16
TOTAL SUPPLY after smallholder storage losses	1.71	1.96	2.20	1.92	2.20	2.48	2.15	2.46	2.77
TOTAL SUPPLY per capita (MT)	0.13	0.13	0.13	0.15	0.15	0.15	0.16	0.16	0.16
TOTAL SUPPLY before losses & estates	1.79	2.08	2.37	2.13	2.46	2.81	2.54	2.93	3.31
TOTAL SUPPLY per capita (MT)	0.14	0.14	0.14	0.16	0.16	0.17	0.19	0.20	0.20
CONSUMPTION									
Total population (million)	13.07	14.99	16.91	13.07	14.99	16.91	13.07	14.99	16.91
Human Consumption (million MT)	1.68	1.93	2.17	1.88	2.16	2.44	2.10	2.41	2.72
Add brewery / animals (million MT)	0	0	0	0	0	0	0	0	0
Total consumption (million MT)	1.71	1.96	2.20	1.92	2.20	2.48	2.15	2.46	2.77
Assumptions:									
smallholder storage losses		15.0%			17.5%			20.0%	
Kg maize/person /day:		0.35			0.40			0.44	
kcal/person/day:		1,800			1,950			2,100	
kcal/kg maize:		3,578			3,578			3,578	
% kcal from maize:		70.0%			72.5%			75.0%	

Adapted from Dorward & Chirwa (2010a). Sources: NSO (2009), Carr (pers comm.), (Jayne et al 2010)

4 Conclusions and recommendations

This paper has examined methodological lessons from experience in estimating yield and incremental production benefits from the Malawi Farm Input Subsidy Programme (FISP). The main conclusions that emerge from this are that

- there is a significant upward bias (or overestimate) in area estimates obtained by relying on information from farmers on the areas of their crops.
- GPS technology appears to provide an affordable but more reliable alternative method for measurement of plot areas,
- estimates of yield from information on total plot harvests are affected by bias in area measurement and, for given estimates of production per plot, over-estimates of area lead to under-estimates of yield.
- bias in farmer estimates of plot production are difficult to judge but may tend to under-estimate yields as a result of
  - unwillingness of farmers to reveal total harvest, and
  - errors in reporting harvests in standard units.
- enumerator harvesting of yield sub plots is costly and without very close supervision and longer engagement with farmers is likely to lead to over-estimates of yield.
- estimates of crop yield response to fertiliser application are over- (under-) estimated if whole plot harvest is over-(under-) estimated by farmers.
- fertiliser responses are also over-estimated if
  - yield sub plots over-estimate yield, especially if plot areas are over-estimated, as this leads to an under-estimate of fertiliser application rate.
  - analysis does not allow for the effects of early planting, plant density, seed type, number of weedings and early weeding (management practices which raise yields but tend to be correlated with fertiliser use).

The following recommendations are made regarding area and yield survey methods:

- all studies should provide a clear description of field and analysis methods used, and of basic descriptive statistics (for example as in table 2.2)
- more studies should attempt to triangulate and compare methods
- IHS2 estimates of crop areas and yields should be treated with caution as they are likely to have over-estimated areas and under-estimated yields
- GPS should normally be used for measuring plot areas but further investigation is needed to establish accuracy and possible bias when measuring small plot areas.
- continued work is needed to develop accurate methods for estimating plot yields – possibilities that could be investigated include
  - intensive measurement of whole plot harvests from a sample of plots to ‘ground truth’ remote sensing yield estimation methods with, for example, spectral analysis of electromagnetic radiation linked to crop yield models
  - further study of the accuracy of farmer reported harvests, with greater refinement of standard crop measures
  - development and testing of alternative yield sub plot measures and systems, and estimation of the determinants of bias
- analysis of subsidy impacts should separate out the effects of these practices and also estimate how far management practices are changed by receipt of subsidised inputs. This may need to rely on formal trials to address problems of multi-collinearity across management practices
- high prevalence of inter-cropping of maize with other crops means that it is important that yields of intercropped maize are properly analysed and reported.
- a review of IHS3 results and experience in addressing the issues raised here should make an important contribution to development of improved survey methodologies, which are themselves critical for agricultural policy planning and implementation

Acronyms

AISS1	Agricultural Input Subsidy Survey 1 (2006/7)
AISS2	Agricultural Input Subsidy Survey 2 (2008/9)
FISP	Farm Input Subsidy Programme
GPS	Global Positioning System
IHS2	Integrated Household Survey 2 (2003/4)
IHS3	Integrated Household Survey 3 (2010)
MOAFS	Ministry of Agriculture and Food Security
NACAL	National Census of Agriculture and Livestock (NSO, 2010).
NSO	National Statistics Office
PVA	Poverty and Vulnerability Assessment. (Malawi Government and World Bank, 2006).

References

Chibwana, C., M. Fisher, C. Jumbe, W. Masters and G. Shively (2010). Measuring the Impacts of Malawi's Farm Input Subsidy Program. Draft for discussion at BASIS AMA CRSP TC meeting.

Dorward, A., E. Chirwa (2010a). *Evaluation of the 2008/9 agricultural input subsidy Malawi: Maize production and market impacts* London, School of Oriental African Studies (SOAS).

Dorward and Chirwa (2010b) *The Farm Input Subsidy Programme (FISP) 2009/10: A review of its implementation and impact.* Draft November 2010

Holden S. and Lunduka R. (2010) *Too poor to be efficient? Impacts of targeted fertiliser subsidy programme in Malawi on farm plot level input use, crop choice and land productivity*. Norwegian University of Life Science, As, Norway.

Jayne, T. S., J. Mangisoni, N. Sitko and J. Ricker-Gilbert (2010). *Malawi's Maize Marketing System*, Michigan State University.

Malawi Government and World Bank (2006). Malawi Poverty and Vulnerability Assessment. Investing in Our Future.

National Statistical Office (2009). *2008 Population and housing census: preliminary report*. Zomba, Malawi, National Statistical Office.

National Statistics Office (2010). *National Census of Agriculture and Livestock (NACAL).* NSO, Zomba