Estimating Crop Production in Development Projects
Methods and Their Limitations

C. D. Poate and Dennis J. Casley

A Technical Supplement to Monitoring and Evaluation of Agriculture and Rural Development Projects

Dennis J. Casley and Denis A. Lury

A World Bank Publication
Contents

1. Introduction 1
2. Measuring Production Changes in the Project Context 3
   The Requirement for Output Measurement 3
   Problems in Measuring Production 5
3. Measurement of Crop Production 9
   Harvesting Total Output 9
   Sampling of Harvest Units 10
   Subplot Crop Cutting 14
   Farmer Estimates of Output 23
4. Area Measurement 26
   Ground Transects 26
   Plot Area Measurement 27
5. General Issues 29
   Intraclass Correlation 29
   Sampling Units and the Calculation of Crop Yields 31
   Reporting Results with Mixed Cropping 31
   Scale of Inquiry 33

APPENDIX: Standardization of Crop Cut Data 34
1. Introduction

This booklet is intended to assist those staff engaged in monitoring and evaluation (M&E) of projects who are required to estimate incremental production of annual food crops. It is one of a series designed to review in some depth technical issues which are not covered in the standard texts or which present special problems in the context of monitoring and evaluation. The series supplements Dennis J. Casley and Denis A. Lury's handbook on monitoring and evaluation (hereafter referred to as the Casley and Lury handbook). To encompass the full spectrum of M&E, the Casley and Lury handbook takes a somewhat broad perspective, highlighting issues but not pursuing them in detail. Readers have pointed up the need for more detailed reviews of selected topics, prompting this booklet and one on sampling issues. A detailed case study on the implementation of an M&E program is already available.

This technical supplement addresses the relative merits and limitations of various techniques for measuring crop production on smallholder farms. The section on one of the most popular techniques, crop cutting, draws on analyses that have not been published elsewhere. The data for the analyses, summarized herein, were made available by the Nigerian Federal Department of Rural Development (FDRD) and were taken from studies conducted by the Agricultural Development Projects in northern Nigeria. Given the scarcity of data reflecting the accuracy of the crop cutting technique, the cooperation of the FDRD has been especially important for this review.

The issues covered in this booklet are relevant to annual crop estimation studies in any small-farmer project. However, the dis-

1. Measurement of production of tree crops such as coffee, oil palm, and coconuts presents a set of problems different from those discussed here. In general, with the cooperation of the project beneficiary, the harvest of such crops can be recorded in standard units (the same units in which they are normally sold) over a period of time.
cussions and the new analyses focus sharply on Africa. This is intentional. The complex cropping systems used in Africa present special problems for measurement and reporting; in the past, reviews of crop yield estimation techniques have drawn heavily on studies in Asia.

Specific recommendations are made even though comparative data on the different techniques are still scanty. Many readers are looking for immediate guidance as to which method to apply and do not have the option of waiting for further studies. We must stress that the choice of method depends on the focus and objective of the survey. Some procedures offer high levels of accuracy but require intensive supervision and thus are appropriate for small samples only. Others are suitable for large samples but can give only approximate indications of levels of production. Our recommendations will not always apply, and readers must make their own judgment. But at least they can make such judgments with an awareness of the problems that may be encountered. One principle holds true for all methods: none will work unless high standards of data collection and survey management are set and maintained.
2. Measuring Production Changes in the Project Context

The appraisal documents of many agriculture projects emphasize the quantities of inputs to be supplied and the changes in output that are expected to result. The role of a monitoring and evaluation unit is to obtain data that reflect what progress the project makes toward meeting the input and output targets and to report on the factors that affect the ability to achieve the incremental outputs desired—factors such as the number of farmers using fertilizer for the first time or as repeat users and the constraints that farmers face, given their own perceptions and attitudes, in applying the technical packages. To managers responsible for project implementation, these considerations are often more important than the level of changes in output, which may take time to detect with any clarity. Moreover, when the activities of a project are broken down into discrete steps, the problems of gathering data to monitor the steps become more manageable and the need for sophisticated survey techniques and analysis is reduced. The priorities for data collection are examined in the Casley and Lury handbook. This technical supplement deals with only one aspect of an M&E work program: the collection of data relating to crop output.

The Requirement for Output Measurement

Agriculture and rural development projects vary considerably in their scope, objectives, and components. Frequently, however, they share a need for measurements of crop production as indicators of performance. Such production indicators are usually required in projects where:

- The intended output is a rise in crop production attributable to a combination of factors such as an expansion of the area under cultivation, a change in cropping patterns, an increase in output per unit area, or an increase in output per unit of labor.
- The expected rise in output is founded on a technological package involving the introduction or expanded use of inputs, a
change in cultivation practices, an improvement in the extension services, or better access to markets.

- A technological package, although proven under research trials, has not been fully tested in the project area, and there is uncertainty about the likely response of farmers and the performance of the package when applied in the particular circumstances of the farmers’ own fields.

There are several types of crop-output-related information which may be required, depending on the particular performance issues of the project, and there are different approaches, as described below, that can be used for obtaining the information.

- **Farming system review.** Only rarely is a project planned from a sound data base. Although a planning data base is the responsibility of the project designers, the M&E unit may need to test the plan’s basic assumptions with regard to cropping pattern, use of inputs, and, possibly, global crop yields. For such a test the use of remote sensing devices to generate details of settlement distribution, livestock concentrations, and cultivated areas should be considered. New techniques of photography from low-flying aircraft are relatively inexpensive. Informal studies based on open-ended farmer interviews may also add much to an understanding of the farm systems.

- **Adoption and output indicators.** Simple indicators of the number of farmers adopting a technique for the first time and of repeat adoption should be based on a large sample so that results can be disaggregated to the smallest administrative unit at which the project operates. Within the same sample, approximate indicators of output in local units may be obtained from the farmer; these may at least give some indication that change is under way.

- **Area and yield surveys.** Surveys of randomized samples of farms or fields are used to obtain annual area and yield estimates. Such surveys usually involve objective measurements because it is assumed that farmers are unable to report their crop yields in the quantitative terms desired or with the level of accuracy required.

- **Technical potential.** New production methods for crops and livestock are tested both under intensive research conditions and
under controlled conditions in farmers' plots. Accurate measurement of crop yield is essential.

Farm management studies. The results of crop potential and adoption studies may generate the demand for a detailed survey of the resource use of the farming household. This type of work may be contracted out to a specialized university or institute which has experience in small-scale, intensive measurement studies.

An assessment of changes in crop production can therefore be based on one or more of several measures: baseline and subsequent time series yields, technical potential of innovations under research and farmer conditions, and adoption indicators of farmer response. When each part of the project activity–project output linkage is considered separately, the survey requirements for each stage become clearer. Treating discretely each element of data collection lowers the burden of inference from any one survey and creates an appropriately modest series of objectives for monitoring and evaluation.

Problems in Measuring Production

Despite the fact that a number of indicators are required to reflect project progress and different survey approaches are required for each, the pervasive tendency among M&E units has been to embark on large-scale resource-use and production surveys. The objectives have been, variously, to quantify annual changes in production in order to compute a stream of benefits as a measure of project success; to quantify patterns of labor utilization, farm and off-farm incomes, and household expenditure to make up for the lack of data available at the project-appraisal stage; or to satisfy a perceived need for a broad baseline study against which to compare later studies in order to measure the level of change.

Given the resources available to a typical M&E unit, the result often is a common sample of several hundred or more holdings to meet all objectives. Both the objectives and the feasibility of meeting them are questionable.

When dealing with farmers who do not keep records and who are unable to express their farm output in precise quantitative terms, M&E units tend to rely on multivisit surveys. That way the farmers
can be questioned about such matters as labor inputs while the information is still fresh in their minds, and the surveyors have an opportunity to apply objective measurement techniques to obtain data on areas and yields. Given this methodology, several hundred holdings are a lot to cope with, and a high-quality product is required to justify the cost. Supervision of the enumeration quality of such multivisit surveys is notoriously difficult. Moreover, the volume of data generated by the lengthy questionnaires and multiple visits becomes considerable; many M&E units have been smothered in their own data. The advent of microcomputers may help in this regard, but the experience to date shows that even with the new technology serious problems remain.

However, if the intention is to measure small changes from year to year or to present findings for each geographical subpopulation, several hundred households is too small a sample on which to base the kind of comparisons and conclusions desired. Sampling issues are dealt with in a companion booklet in this series, but the fundamental problems are worth illustrating in the specific context of crop production estimates.

A typical coefficient of variation (standard deviation divided by the mean) of plot yields is 50 percent. To estimate a mean yield with a high level of confidence that it will be within 10 percent of the true mean requires a sample of approximately 100. This calculation assumes a simple random sample. But the sample for production surveys is usually clustered, with one enumerator covering ten to twenty holdings in one settlement or village. The loss of sampling efficiency with a clustered sample (see section 5) is such that for sound yield estimation the size of the sample may have to be increased by a factor of two or three. If the yield estimates are required for various subgroupings in the project area (geographical or project beneficiary versus nonbeneficiary, for example), the estimated minimum sample size is required for each such grouping. And since a given sample holding may not contain the specific crop under study, larger samples may be required. For the estimate of, say, maize yields to meet the above standards when maize is grown

5. This is obtained by using the well-known formula (defined and demonstrated in the Casley and Lury handbook): $n = k^2V^2/D^2$. 
on only 50 percent of the holdings, a further doubling of the sample in terms of selected holdings is necessary.

Estimating the mean yield for a given season is one issue, measuring the underlying trend of production over a span of years presents another problem. Factors outside the control of the project, such as climate and price policy, will introduce variations in the time series that may be greater than the long-term trend which, it is hoped, is at least partly induced by the project stimulus. The number of time points required to estimate the slope of a linear time series is a function of the random variation from one time point to another and the required precision of the estimate. If the slope is 10 percent of base yield and it must be estimated with a standard deviation of 2 percent of base yield, and assuming a random variation of 15 percent owing to exogenous factors, it can be shown from the formula for the variance of a regression coefficient that approximately nine time points would be required. If the annual increment being estimated is 100 kilograms per hectare, the confidence interval of the estimate with nine time points would be approximately 60–140 kilograms. With four to five time points, it will be difficult to detect a yield trend that is rising even at 10 percent a year and be sure that it is significantly higher than zero.

But a nine-point time series is far in excess of the life span of most projects. In many projects, multivisit surveys have been maintained for four or five seasons at most.

It must be stressed, therefore, that the determination of yield trends is a dubious M&E objective unless the exercise will be continued well past project completion.

The argument that because of sample size and time span requirements the objective of measuring crop production changes may be infeasible is reinforced by the need to consider the question of bias in the data collection methods. The overall mean square error of an estimate is a function of both the sampling error and the nonsampling error, the latter including data collection biases. As will be seen below, the techniques in common use for measuring crop output are subject to high levels of bias unless great efforts are made to supervise their execution.

All this having been said, the demand for key indicators of production performance is a real one. It is important, therefore, to consider what can be done to provide at least some indication of
production movements for purposes of project evaluation. (From the argument above, it is obvious that managers cannot rely on obtaining production trends as a monitoring indicator of project implementation.) For example, a rough estimate by a sample of farmers widely dispersed across the project area may provide an indication of "change under way," even if precise quantification is difficult. And if sampling requirements are relaxed, high-quality case studies on purposively chosen adopters and nonadopters of a project stimulus may provide insight into the performance of the stimulus.

Our principal objective here is to demonstrate that the choice of the method of measuring crop production depends on the type of survey to be undertaken. For example, crop cutting can produce results of reasonable accuracy, but only if the fieldwork is closely supervised. Crop cutting is more suitable for a detailed study of crop response than for project-wide estimation of crop output. Conversely, under certain circumstances, farmers' estimates of their crop output, which can be obtained from a large sample, will be no more biased than crop cutting on a sample of similar size and can be collected without great expenditure of resources and skills.

Given that the more objective methods are suitable for small-scale studies, the worst mistake, and a common one, is to attempt a large random sample coverage (to provide the level of disaggregation sought) using the methods appropriate for small samples and using junior and inexperienced staff.

Fit the survey size to the method and vice versa is the common principle underlying the recommendations in this booklet.
3. Measurement of Crop Production

Most techniques for measuring crop yields involve the sampling of crop output; total crop output is estimated from the harvest of a subplot of land or by weighing samples drawn in local harvest units from the aggregate harvest of the farming household. Later in this section we will focus on a number of issues germane to all measurement techniques which derive estimates of production from samples. Before turning to the sampling methods, however, we first consider the most straightforward measure of output: direct weighing of the total harvest. For purposes of our discussion the following definitions are employed:

*Holding*—the total area farmed by a household (including fallow land). A holding may include land owned or rented by the household.

*Field*—a contiguous piece of land which the farmer considers to be a single entity.

*Plot*—a subdivision of a field, which subdivision contains a single crop or a homogeneous mixture of crops. In most cases a plot is defined to meet the needs of the surveyor; the boundary definition may not be recognized by, or meaningful to, the farmer.

*Subplot*—an area of land within a plot, which area is marked by the surveyor for the purpose of crop cutting or other measurement operations.

*Harvesting Total Output*

The idea of harvesting a complete plot is usually rejected because of the volume of work required, especially if all the crops within a

6. Other issues, although important, lie outside the scope of our discussions. For example, whichever method is used to obtain harvested yields, the survey should include measurement of the moisture content and threshing percentages for each crop. The treatment of missing data and low or zero yields should be considered, particularly if the low yields are a result of damage by wildlife or other natural phenomena.
holding are to be recorded. Given a range of crops and even a moderate-size sample, the weighing of the total harvest is clearly not a viable option. However, there are circumstances in which total harvesting would be suitable, and the method has several advantages over any of the sampling methods, whose results are likely to be biased, often seriously so. With crop cutting, there is the further danger that what will end up being measured will be the biological yield rather than the economic yield, since the subplot is likely to be harvested more thoroughly than the main plot. Yet it is the economic yield that is generally of greater interest to the project.

If an M&E unit is conducting a case study of a particular smallholder crop or farming practice, the weighing of a total harvest is an appropriate and feasible option. Since the farmer has to harvest the plot at some time anyway, nothing extra is required of him other than his cooperation in allowing the enumerator to weigh the produce. Should the enumerator arrive a day or two late, the farmer would have to make sure that the crop is retained intact. In studies involving project beneficiaries it should be possible to obtain this level of cooperation. Project staff are not total strangers to such beneficiaries—or should not be. Given such cooperation, the method may be particularly appropriate for cash crops that are retained on the holding for a brief period prior to sale.

In a situation where there is as much variation in crop yield within plots as there is between plots (see discussion below under "Number of subplots"), a complete harvest is advisable if the intention is to estimate a crop response function using the yield data as the dependent variable.

It is strongly recommended that for microstudies on the typically small plots of smallholder beneficiaries serious consideration be given to measurement of the complete harvest of the farmers’ plots.

**Sampling of Harvest Units**

A method of measuring crop production that requires neither a total harvest nor a crop cut involves sampling of the farmer’s harvest after it has been gathered in and before it is transported to the house or store. At harvest time the farmer is left to collect the produce in his normal harvest units, such as sacks, baskets, bowls, or bundles of grain heads, or individual roots and tubers. The enumerator visits
the plot and inspects these units. A subset is sampled and weighed. Total plot output is obtained by calculating the mean harvest unit weight and multiplying it by the number of units harvested. Yield is obtained by dividing the total output by the area of the plot. (For observations on measuring plot area, see section 4.)

Although this technique is sound in concept, a number of practical issues affect its operation. The crux of the method is the estimation of mean harvest unit weight. For this estimation to be valid, the harvest units must be the correct and complete units for the plot being studied and the sample must be unbiased. Unit weight is a difficult measure to check. The uncertain timing of harvests complicates routine supervision of the enumerator. Furthermore, with many types of crop the units are removed from the plot and transported to the store or otherwise disposed of in short order; once that happens, the original group of units cannot be resampled. Experience with the use of this method in northern Nigeria indicated that many enumerators simplified their task by weighing only one or two bundles (the local units for harvesting sorghum and millet) and imputing weights to the others. Table 1 shows the variation in bundle weights for two crops at three locations over several years. Although each of the three sites shows a measure of internal consistency over the years, there are large variations between sites and in the coefficients of variation between years.

The estimation of harvest unit weight is further complicated by the second weakness of this technique, that of plot identification. Recall the definition of a plot—a subdivision of a field, which subdivision contains a single crop or a homogeneous mixture of crops. Although many users of data require crop areas and yields to be classified in this way, the definition may have no meaning for the farmer who thinks of a field as one entity.

By way of example, consider a field planted to a short-season millet crop at the onset of the rains. The surveyor would define this as one plot occupying the whole field. If at a later date sorghum is planted on, say, half of the field, the surveyor would redefine that field as containing two plots: plot 1 planted to a sole crop of millet and plot 2 planted to a mixture of millet and sorghum. Sorghum is a long-season crop, so the millet is harvested while the sorghum is still growing. To the farmer the millet in the field is one entity, to be treated in the same way whether it is harvested from the sole-crop or
Table 1. Variation in Bundle Weights

<table>
<thead>
<tr>
<th>Item</th>
<th>Funtua</th>
<th></th>
<th>Gusau</th>
<th></th>
<th>Gombe</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sorghum bundles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (kilograms)</td>
<td>46</td>
<td>49</td>
<td>26</td>
<td>30</td>
<td>32</td>
<td>31</td>
</tr>
<tr>
<td>Coefficient of variation (percent)</td>
<td>50</td>
<td>n.a.</td>
<td>n.a.</td>
<td>23</td>
<td>59</td>
<td>18</td>
</tr>
<tr>
<td>Number of plots</td>
<td>445</td>
<td>375</td>
<td>n.a.</td>
<td>400</td>
<td>462</td>
<td>55</td>
</tr>
<tr>
<td>Millet bundles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (kilograms)</td>
<td>42</td>
<td>31</td>
<td>26</td>
<td>29</td>
<td>31</td>
<td>n.a.</td>
</tr>
<tr>
<td>Coefficient of variation (percent)</td>
<td>57</td>
<td>77</td>
<td>n.a.</td>
<td>21</td>
<td>52</td>
<td>n.a.</td>
</tr>
<tr>
<td>Number of plots</td>
<td>264</td>
<td>206</td>
<td>n.a.</td>
<td>347</td>
<td>346</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

n.a. Not available.

the mixed-crop portion. To the surveyor who seeks a plot-specific estimate of output, it is important that the millet from the two plots not be mixed. Unless the enumerator is physically present during the harvest operation, the chances of the two outputs being correctly identified are slim. This example is a simple one. Further complications can easily arise. For example, there may be a third crop. Or consider the reverse of the example: if the millet were planted on just half the field and the sorghum on the whole field, there would be no clear demarcation of the plot boundary by the time the sorghum was ready for harvest.

The timing of the harvest measurement can be a complex issue if the field is large and the harvesting operation is spread out over several days or longer, or if the crop requires multiple harvesting. Cotton, yam, cassava, and many vegetables may not be harvested in one operation but over a period of several weeks. In such circumstances the enumerator cannot always be present for the whole harvest in the field. Not only does this exacerbate the problem of assigning units to plots, but the sampling of units becomes a haphazard process. Multiple harvesting also occurs with such crops as maize, which may be partially harvested for immediate consumption while still green.

Root crops present a special problem in addition to that of continuous harvesting. In the case of yam and cassava, there is often no traditional harvest unit. Because of their bulk the tubers are handled individually. The process of estimating mean tuber weight and counting the total number of tubers from multiple harvests is fraught with potential for error. The harvest unit method is not recommended for yam and cassava. Its utility with regard to crops of a smaller total bulk, such as cocoyam or potatoes, must be judged within the particular circumstances of the survey.

To obtain a satisfactory estimate of crop yield at the plot level requires both consistency and accuracy in three separate estimations: first, the area of the plot; second, the number of units harvested; third, the average unit weight (under standard crop conditions). A mismatch or error in any of these items will invalidate the final estimate. Evidence suggests that if the harvest unit technique is used on a large sample, the plot estimates will be unreliable. If this method is used, production should be expressed on a per holding
basis only, to avoid confusion regarding the matching of units to plots.

In general, therefore, it is recommended that the measurement technique of sampling harvest units to obtain a mean unit weight be used only when the production estimate per crop is required on a per holding rather than per plot basis.

**Subplot Crop Cutting**

Although crop cutting is in widespread use as a survey technique, relatively little written guidance is available for field survey units. It is widely recognized that this method results in overestimates of yield. Casley and Lury attribute the overestimation to a combination of edge effects (plants that lie fractionally outside the subplot are included in the count), border bias (location methods may over- or underrepresent the boundaries of a plot), and nonrandom location of the subplot (enumerators tend to avoid bare or sparsely populated parts of the plot). Moreover, in effect it is biological yield that is measured; careful estimation of crop losses is necessary if this is to be accurately converted to economic yield. Our review of the crop cut method of yield estimation focuses on four major issues: the size of the subplot, the number of cuts to be taken, the shape of the subplot, and the location of the subplot.

**Size of the Subplot.** The question of what size of subplot to use for crop cutting purposes, given the tradeoff between accuracy and ease of operation implicit in such a choice, has stimulated more investigation than any other aspect of the crop cut technique. Pioneering studies by P. C. Mahalanobis on jute and paddy rice and by P. V. Sukhatme on rice and wheat remain the authoritative works on this topic. The most often quoted results are reproduced in table 2. They lend substantial weight to the proposition that given a subplot of moderate size, the percentage of overestimate is low; thus the technique offers a practical and objective approach to crop yield estimation. Indeed, it may be inferred from the table that, for all practical purposes, unbiased yield estimates are possible with subplots larger than 40 square meters. The experiments in wheat found no overestimate at this level. Despite doubts about the general validity of this conclusion, given the limited range of crop types
Table 2. Overestimation of Yield with Small Plots

<table>
<thead>
<tr>
<th>Type of crop and shape of subplot</th>
<th>Size of subplot (square meters)</th>
<th>Number of plots</th>
<th>Overestimate (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wheat—irrigated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equilateral triangle</td>
<td>43.80</td>
<td>78</td>
<td>0.0</td>
</tr>
<tr>
<td>Equilateral triangle</td>
<td>10.95</td>
<td>78</td>
<td>4.8</td>
</tr>
<tr>
<td>Equilateral triangle</td>
<td>2.74</td>
<td>78</td>
<td>15.7</td>
</tr>
<tr>
<td>Circle</td>
<td>2.63</td>
<td>117</td>
<td>14.9</td>
</tr>
<tr>
<td>Circle</td>
<td>1.17</td>
<td>117</td>
<td>42.4</td>
</tr>
<tr>
<td>Wheat—unirrigated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equilateral triangle</td>
<td>43.80</td>
<td>107</td>
<td>0.0</td>
</tr>
<tr>
<td>Equilateral triangle</td>
<td>10.95</td>
<td>107</td>
<td>11.0</td>
</tr>
<tr>
<td>Equilateral triangle</td>
<td>2.74</td>
<td>107</td>
<td>23.4</td>
</tr>
<tr>
<td>Circle</td>
<td>2.63</td>
<td>162</td>
<td>14.8</td>
</tr>
<tr>
<td>Circle</td>
<td>1.17</td>
<td>161</td>
<td>42.4</td>
</tr>
<tr>
<td>Paddy rice</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rectangle</td>
<td>40.47</td>
<td>108</td>
<td>0.8</td>
</tr>
<tr>
<td>Circle</td>
<td>2.63</td>
<td>216</td>
<td>4.5</td>
</tr>
<tr>
<td>Circle</td>
<td>1.17</td>
<td>216</td>
<td>9.0</td>
</tr>
</tbody>
</table>


included in the Indian studies from which the data were taken, crop cutting has become a popular method of yield estimation and has been used with a great variety of subplot sizes.

The most recent study by the Food and Agriculture Organization concludes that the size of subplot used in crop cutting should be "a function of the density of the crop within the field. For the very dense, irrigated . . . [crops] . . . the plot size could be quite small: 1–5m². For more widely spaced crops like maize, tubers, etc., the plot size could be larger: 10–25m². While, for very widely spaced crops and in the case of mixed cropping, the plot size could be as large as 100m²."  

This vagueness is reflected in the somewhat arbitrary combinations of size and number of subplots which have been used in attempts to control bias. In Nigeria, crop cutting was adopted in 1980 as the method of yield estimation in all M&E surveys. To test the

applicability of the Indian findings to conditions in the West African savannas, a crop cutting study was carried out in yam and sorghum fields. A total harvest was included to provide the actual yield against which to compare the yields obtained from crop cuts of different sizes. Combinations of one or more 50-square-meter triangles and 100-square-meter squares were studied, and it was found that the overestimate could be as high as 28 percent for the sorghum crop and 17 percent for the yam when only one 50-square-meter triangle per plot was used. In particular, the study concluded that there was little improvement to be gained from sampling more than one subplot unless the area sampled increased to 200 square meters.\(^8\)

The error measured in this and other studies is a combination of the bias inherent in the technique and the sampling error of the subplot observations.\(^9\) The sampling error can be reduced by increasing the number of subplots laid. The main issue, therefore, is the magnitude of the bias and its response to increasing subplot size. For a more rigorous analysis than was originally the case, the data from the Nigerian yam and sorghum crop cutting study were reanalyzed as described in the appendix. The analysis revealed that the mean overestimate using a 50-square-meter triangular subplot was in the range of 10–14 percent for both sorghum and yam. And the bias is not reduced as the size of the subplot is increased above 50 square meters.

Since these studies were conducted under careful supervision, the size of the bias is substantially larger than one would have wished. The consistency of overestimation, however, offers some hope that the crop cut method could be used if the results are adjusted for the upward bias.

**Number of subplots.** Where the crop under study is grown in evenly planted, dense stands and under carefully controlled conditions, the level of within-plot variation would be expected to be low

---

\(^8\) These were the results reported internally within Agricultural Projects Monitoring, Evaluation, and Planning Unit (APMEPU), the unit responsible for the surveys. In the event, for the evaluation surveys of the agricultural development projects, the subplot size was standardized at a 100-square-meter right-angled triangle.

\(^9\) Of course, the sampling error may not be in the same direction as the bias in any particular sample.
compared with the variation between plots. Analysis of the Nigerian data used above reveals a high level of within-plot variation; this finding has been confirmed by an independent set of observations from Niger. The results are reported in table 3, where the variation within plots is seen to be at least 40 percent of the total variation for all of the crops surveyed and to exceed 50 percent in the case of yam.

In view of these levels of within-plot variation, the question of the number of subplots to be sampled within a plot needs consideration. The sampling error decreases in proportion to the square root of the number of subplots laid. The standard error of a within-plot estimate from two subplots is approximately 70 percent of the magnitude of the error from one subplot. With three subplots the standard error decreases to about 58 percent of the error from one subplot.

Given that two or three subplots would appear to be necessary if reasonably precise estimates of plot yields are required, the question of enumerator workload arises. More work is involved in harvesting two 50-square-meter subplots than in harvesting one 100-square-meter subplot because of the extra effort required to locate, lay, and demarcate the subplot boundaries. One might think that the objective of laying more than one subplot could be met by laying the subplots in a contiguous pattern; thus two 50-square-meter triangles could be combined into one 100-square-meter square. Unfortunately, this argument fails because of the form of the within-plot variation. The pattern of crop yield variation within a field resembles a clustered rather than a random distribution. Thus two contiguous subplots will not, in practice, be independent. This can be illustrated by comparing deviations of subplot yields about their mean for sets of two separated and two contiguous samples drawn from the Nigerian sorghum data supplied by APMEPU: the mean deviation of the separated subplots was 240 kilograms; the mean deviation of the contiguous subplots was 149 kilograms. Since deviations between separated subplots are some 60 percent greater than those between contiguous subplots, independently located subplots are required to improve the within-plot estimate of crop output.

The particular combination of size and number of subplots to be used depends on the particular circumstances of the survey. There
Table 3. Variation between and within Plots

<table>
<thead>
<tr>
<th>Country and crop</th>
<th>Subplot size (square meters)</th>
<th>Number of subplots per plot</th>
<th>Number of plots</th>
<th>Variation between plots (percent)</th>
<th>Variation within plots (percent)</th>
<th>Total variation (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nigeria</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sorghum</td>
<td>50</td>
<td>6</td>
<td>30</td>
<td>55</td>
<td>45</td>
<td>100</td>
</tr>
<tr>
<td>Yam</td>
<td>50</td>
<td>6</td>
<td>31</td>
<td>42</td>
<td>58</td>
<td>100</td>
</tr>
<tr>
<td>Niger</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Millet (1982)</td>
<td>30</td>
<td>3</td>
<td>99</td>
<td>60</td>
<td>40</td>
<td>100</td>
</tr>
<tr>
<td>Millet (1983)</td>
<td>30</td>
<td>3</td>
<td>103</td>
<td>52</td>
<td>48</td>
<td>100</td>
</tr>
</tbody>
</table>

Sources: Data for Nigeria: APMEPU, FERD, Nigeria; data for Niger supplied by J. McIntire, International Crop Research Institute for the Semi-Arid Tropics (ICRISAT), Niger.
is no benefit to be gained from using a subplot larger than 50 square meters, but the accuracy of the plot estimates will improve if more than one subplot is laid. A combination of three 30-square-meter subplots was used successfully in the Niger study reported in Table 3. Alternatively, two 50-square-meter subplots may be adequate and would involve somewhat less demarcation work.

The recommended configuration for the unevenly planted, low-density, irregularly shaped plots common to many African countries is either two 50-square-meter or three 30-square-meter subplots. The only exception to this would be in the relatively uncommon situation where the objective of the survey is to estimate the mean yield of the overall project area rather than to perform any analysis based on characteristics which are specific to the plot. In this situation the recommended approach would be to lay only one subplot per plot; this would ensure maximization of the number of plots sampled for a given enumerator resource and a better dispersion of the sample over the target population. This type of survey objective is normally associated with national statistical studies; it will rarely be appropriate for an M&E program whose objectives include the measurement of changes in output and the attribution of causality.

There are, however, other operational factors to be considered when more than one subplot is to be laid out. The gains in accuracy to be derived from laying two subplots will not be achieved unless the subplots are statistically independent of each other. The data for both the Nigerian and the Niger studies quoted above were collected under closely supervised conditions that cannot normally be replicated when a large sample is surveyed. Experience has shown that without this close supervision the subplots will not pass the test of statistical independence. Enumerators tend either to locate the subplots in areas where high yields are expected or, worse, to create fictitious data after harvesting just one subplot.

Two examples from recent experience in Africa illustrate such problems of enumerator bias. In a closely supervised experiment encompassing seven villages, enumerators estimated plot yields of millet and sorghum by crop cutting two subplots per plot; they also weighed the total harvest from each plot. In four of the seven villages the coefficients of variation recorded from the subplots were only half the order of magnitude of the coefficients of variation from
the harvest of the corresponding whole plot. The subplots were, in other words, much less variable than the whole plots—an implausible result. The correlation coefficient of the differences between subplot yields and the corresponding whole-plot yields was extremely high: greater than 0.8 in eleven of fourteen cases. If one subplot overestimated the whole-plot yield by a large amount, the second, supposedly independently located subplot usually overestimated by the same order of magnitude. Inspection of the records revealed enumerator-specific cases of pairs of subplot yields which were implausibly similar in size and direction of bias. In a second example two subplots per plot were laid as part of the procedure in a project survey. The coefficients of variation for the major crop were higher than in the previous example but were still less than 40 percent of the mean yield in eight out of thirteen villages. The level of correlation between subplots was greater than 0.7 in six of the thirteen villages.

Although these surveys were conducted under controlled conditions, the evidence indicates that many enumerators did not follow the rules necessary to ensure independence of the subplots.

If the recommendation of laying two or three subplots is followed, two tests should be applied to the data: Is the correlation coefficient between subplots greater than 0.7, and is the coefficient of variation below 40 percent of the mean yield? If the subplots are statistically independent, the expected correlation between subplots is zero. In practice, because whole-plot yields vary widely, some correlation will be present because the subplots will be correlated with their whole plots. However, since variation between plots is of the order of 50 percent of total variation, it may be argued that a correlation coefficient greater than 0.7 (equivalent to a coefficient of determination of 0.5) would not be expected with independent subplots. A low coefficient of variation is another indication that rigorous randomization may not have been achieved in the subplot location. Samples of crop cuts commonly display coefficients of variation of the order of 50 percent of the mean or more.

If either of the tests is positive, a careful review of the data is indicated. Because nonindependence of subplots is an enumerator-specific error, the tests should be carried out for each enumerator.

Shape of the Subplot. The shape of the subplot can give rise to bias in two ways: First, the shape may be distorted as the boundaries
of the subplot are laid out; this would lead to a change in the dimensions, and thus the area, of the subplot. Second, the boundary lines may be laid out in such manner that the status of strands of crop growing along the boundaries is uncertain—should the strands be included in, or excluded from, the crop cut? This second problem seems to imply that the ideal plot is one which minimizes the subplot perimeter for a given area—that is, a circle. However, the practical problems of demarcating and manipulating anything other than a small circle are considerable, so the choice is normally between a square and a triangle.

In a study by Mahalanobis and Sengupta a square was found to overestimate a circle by 3.5 percent and a triangle to overestimate the circle by 23.5 percent. Although the magnitude by which the bias of the triangle exceeds that of the square is higher than might be expected, the results are not inconsistent; a square has a smaller perimeter than a triangle for any given area. The square, therefore, is the recommended shape.

The orientation of the subplot is another important consideration. With crops planted in rows, the square should be laid with a diagonal along the center of the furrow between rows. If a side of the square is laid parallel to the crop rows, it is possible that the status of a row could be in doubt. And if one side lies along a row, there could be uncertainty over whether to include the plants in that row. When the subplot is laid with a diagonal parallel to the rows, the area of uncertainty is confined to the opposite corners of the square (see figure).
The square is laid in the field with the aid of a single length of rope. The rope should be thick enough so that it will not stretch when wet. To minimize the bulk of the equipment to be carried, the rope is cut to a length equal to the lengths of two sides of the square plus one diagonal. Knots or ties are used to denote the side corners. The square is formed from two triangles with a common diagonal. After the corners have been pegged, the length of each of the diagonals should be checked to ensure that the shape is a perfect square.

Location of the subplot. Because most smallholder plots are irregular in size and shape, the subplot must be located at random within the plot. Two location schemes are in common use. The first requires the enumerator to choose, by eye, the principal diagonal of the plot. He then walks along the diagonal and stops at approximately equal intervals, corresponding to the number of subplots to be laid. This scheme gives the enumerator considerable leeway in positioning the subplot. In view of the evidence that even the most diligent enumerators are inclined to bias the location toward areas of apparently high yield, this procedure is not recommended.

The alternative is for the enumerator to use a pair of random numbers as coordinates. Using random-number tables, the enumerator selects two numbers that lie between zero and a value equal to half the perimeter of the plot. The first number prescribes the distance to be paced around the perimeter from a given point A. The second is the distance into the plot from the entry point given by the first number. If the second number is greater than the width of the plot, the enumerator reverses his path and continues pacing back across the plot until the selected number is reached. The subplot is laid from this point.

On occasion the coordinates selected will be such that the subplot will lie across the boundary of the plot and part of the subplot will fall outside the plot. There are three rules that can be followed in relocating the subplot. Each one introduces bias based on varying probabilities of including the border of the plot.

1. If the method in use is that the subplot is laid forward from the location point, when the location crosses the plot boundary, the lie is reversed so that the plot is laid back along the path
MEASUREMENT OF CROP PRODUCTION

taken by the enumerator. The effect of this rule is to give a high probability of location to a zone parallel to the perimeter of the plot and (in the case of a square subplot) a distance inside the perimeter equal to two diagonals of the subplot.

2. A better procedure, and one that results in less bias, is to slide the subplot back into the plot. The effect of this is to create a narrow zone near the perimeter with a low probability, which is balanced by a higher probability for the remaining area up to one diagonal inside the perimeter.

3. The third option is to reject the location and start over again, selecting another pair of random numbers. This gives the border area an overall lower probability of selection, but this bias is lower in magnitude than that associated with either of the alternative rules.

Although the third choice may, in theory, minimize the bias, it involves extra work, and many enumerators may prefer to modify their pacing routes rather than undertake a relocation. This danger emphasizes the need for supervision of the subplot location operation. The random number coordinates should be recorded on the survey form, and field supervisors should use them to check the accuracy of the subplot location.

Farmer Estimates of Output

The fourth method of crop production estimation—farmer estimates of output—differs from the whole plot harvest, standard unit weight, or crop cut methods in that no direct, objective measure is required for each selected farmer. In certain well-defined cropping situations, carefully obtained farmer estimates can provide valid indications of the year-to-year changes in production for approximate macro-level overviews.

The first requirement for this method is the most limiting. The method can be used only when the farmer collects the harvested crops in units, either traditional or modern, which are consistent and more or less standardized. A good example is the hessian sack designed to hold 50 kilograms or 100 kilograms of grain. Other units exist. The data in table 1 report sorghum and millet production in
Nigeria in bundles, the harvest unit traditionally used in Nigeria for those crops.

If such a harvest unit does exist, the mean weight per unit must be estimated. In view of the variation apparent from table 1, it is recommended that, where possible, a mean unit weight be determined for every village or group of villages from which a small-holder sample is drawn. The units selected for the purpose of determining the mean weight should be inspected to ensure that they do not contain contents other than the crop in question in the appropriate condition.

The sample farmers are asked to estimate their production in terms of number of units. This can be either preharvest (the expected output) or postharvest. Each demands a different approach for enumeration. The preharvest estimate is best done plot by plot, with the enumerator and the farmer in visual contact with the growing crop. This way the enumerator can judge the validity of the response and probe for inconsistencies in the farmer's estimate. After harvest, however, the estimate should be made at the farmer's house so that, if necessary, the enumerator can refer to the farmer's storage capacity as a simple cross-check. In either case, to avoid the need for a precise definition of a plot, the results should be expressed only as an aggregate for the holding.

In one of the Nigerian surveys designed to test the accuracy of the crop cut method, the farmers were asked to estimate the plot yield before either the subplots or the whole plots were harvested. The farmers reported their yields in bundles. A single mean bundle weight was used for the entire sample. The bias in the farmer estimates was estimated in the same way as that for the crop cut method discussed earlier.

The results indicated a bias in the farmer estimate method of approximately 14 percent—the same order of magnitude as that for the crop cut method. This is an important finding since the farmer

10. Studies in Thailand and the Philippines also offer some evidence that farmer estimates result in mean yields that are not substantially different from those obtained from crop cuts (in these studies the result of the crop cut was taken as the standard). In a recent World Bank-supported research study on the impact of extension in Haryana state in India, farmers' statements of both production and crop inputs were used in production functions that explained more than 90 percent of the variation. Those responsible for this study consider that the farmer estimates were of a satisfactory order of accuracy.
estimate method is much simpler, permits a larger sample, and avoids the need for a heavily clustered sample. With farmer estimates, however, it is unlikely that the bias can be reduced appreciably by reducing the scale of the survey and increasing the level of supervision; the crop cut technique does offer this possibility.

In conclusion, then, if complete harvesting is impractical, a carefully executed and closely supervised crop cut is the appropriate technique for small-scale studies that seek to analyze yield as a function of input variables. For broad indicators of holding-level production variations, however, farmer estimates—used in conjunction with estimated unit weights and taken from a sample of appropriate size—are the proper approach. With the latter method it is recommended that the size and direction of bias in farmer reporting be calibrated for the particular circumstances of the project. This requires that a trial be conducted prior to the crop production survey.  

Large surveys that aim to measure production changes for a major project area, by whatever method, are unlikely to determine trends with any confidence for some years. Year-by-year indicators are useful, but they are indicative only and will not permit the calculation of changes in crop production for an economic analysis.

11. Given the shortage of data on the accuracy of smallholder estimates, readers who engage in such a calibration study are urged to record the results for wider dissemination. The M&E unit in the World Bank would be pleased to receive any such comparative data.
4. Area Measurement

If one of the methods described in section 3 is used to obtain estimates of crop production per holding, and the number of holdings is known, it may not be necessary to estimate the area under a crop. But in many cases estimates of crop areas are required, either because expansion of the cultivated area is itself a project objective or because the crop cut method results in a yield per hectare which then must be multiplied by the area to obtain production estimates.

The techniques for measuring plot areas are better known than the procedures for estimating crop yield and are merely summarized here. First, however, we consider the ground transect, a technique for obtaining approximate area estimates by cropping pattern. Ground transects are now somewhat out of fashion, but they may be useful in the context of an M&E survey.

**Ground Transects**

In most circumstances the farm holding is the unit of study for M&E surveys of crop production. But for some purposes, such as a rapid assessment study of a new project area, an overall estimate of the spatial distribution of crops may suffice. The ground transect is a valid method in such a situation. It can be carried out by a relatively small team operating on a mobile basis, it can be done quickly, and it does not require a holding sampling frame.

The technique involves dividing the study area into a grid framework derived from maps or aerial photographs. The choice of grid size is dependent on the size of the study area, the degree of homogeneity in land use patterns, and the resources of the survey organization. Grid squares of 10 kilometers by 10 kilometers have been used successfully. The procedure is to locate a predetermined point in each square (such as the center) and then walk along a randomly selected compass bearing for a fixed distance, say 1 kilometer. Along this transect, observations concerning land use, cropping pattern, and so on, are taken at regular intervals, such as every 20 meters. Summation of the number of points falling on a
specific crop or land use feature expressed as a proportion of the total number of points in the grid transect gives the proportion of land falling in that category.

To estimate the sampling error, a second transect is conducted in the same grid square. This transect may be parallel to the first one, or on a new bearing from the starting point, or independently selected. Stratification can be used to increase grid density in highly cultivated areas and to reduce grid density in sparsely cultivated areas.

There are problems in using this technique. Timing is critical if the cropping pattern includes early- or late-season crops which overlap with the main crops for only a short part of the growing season. A recent study in Kaduna state in Nigeria was carried out in September-October; as a result the area of the short-season early millet crop, which is harvested in September, was underestimated. Physical mobility may be a problem in difficult terrain such as swamps or hilly land. Although the technique is simple, the field staff must be able to interpret a map or an aerial photograph to locate the starting point.

Plot Area Measurement

There are two satisfactory methods of measuring the size of farmers’ plots in order to calculate crop areas and production per unit area. The more popular is the tape and compass, but the dumpy level is also quick and accurate under slightly restrictive conditions.

Dumpy Level. The dumpy level is a surveyor’s level with a telescope fixed to a horizontal base. The telescope may be traversed horizontally and the angular traverse measured with great accuracy. The level is set up near the middle of the plot to be surveyed. An assistant stands with a ranging rod on the plot perimeter at the corner of two sides. The angle of the telescope is zeroed in to this initial point. The rod is then moved clockwise to each corner in turn. The telescope bearing is noted, together with the distance from the dumpy level to the corner, by means of a range finder on the dumpy level. The procedure is repeated until the assistant returns to the initial corner. The surveyor then moves the dumpy level a few meters from its existing position, and a new set of readings is taken.
The plot area is calculated by summing the area of each triangle formed by the angle of traverse between successive corners of the plot and averaging the two estimates for each plot. The area may be plotted on graph paper, calculated by hand using simple trigonometry, or calculated with a simple program on a calculator.

The dumpy level method is quick and accurate, but it is suited for use with low-growing crops only, because a clear line of sight from the center of the plot to the perimeter is essential. (With tall crops, measurements can be taken only when the crops are immature or after they have been harvested.)

TAPE AND COMPASS. The second method of plot area measurement utilizes a measuring wheel or tape and a hand-held compass. A rough sketch of the perimeter of the plot is drawn, and each corner is marked with a letter, starting with A. Each side—AB, BC, CD, and so on—is measured and surveyed in turn. The distance from A to B is recorded in meters. The magnetic bearing with respect to north is recorded from A to B, and then a back bearing is taken from B to A. This information is recorded for each side until the surveying team returns to point A.

The plot area is calculated either by plotting the survey information on graph paper or by using an algorithm on a programmable calculator. When the area is plotted, the final point A (called A’) may not be coterminal with the initial point A. The distance between A’ and A, expressed as a percentage of the measured perimeter A-A’, is termed the closing error. If small, this error can be accommodated as part of the area calculation process. Large errors (those greater than 3 percent) usually call for a resurvey.

12. Other devices such as measuring chains, range finders, and pedometers have been used with some success.

13. FAO. “The Estimation of Crop Areas and Yields in Agricultural Statistics.”
5. General Issues

*Intraclass Correlation*

In practice, the choice of sample design and sample size is dictated by the amount of resources available for construction of a sample frame and by the number of enumerators available to carry out the survey. In particular, with surveys of rural households, a compromise must be reached between the total number of households which can be covered by one enumerator and the physical distance within which an enumerator can move and still keep in close contact with the sampled households. Yield measurement by one of the objective methods requires the enumerator to forge a close working relationship with the farmer. Good cooperation is also necessary when farmer estimates of yield are used, but in such cases the enumerator does not need to be forewarned of impending harvests, nor does he need to become familiar with the farmers' plots. Whole-plot harvesting will be used only for small samples in closely controlled microstudies. The concern, therefore, is with either crop cutting or the unit weight method, when used on large samples.

Crop production surveys are usually based on a multistage design. With two stages the primary units will be settlements or groups of settlements and the secondary units will be farm holdings. More stages may be introduced. The use of settlements or settlement groups on a geographical basis as a primary unit is a form of cluster sampling. This approach reduces the work load involved in constructing a sampling frame and benefits the enumerator in that a set of respondents will be relatively close together. However, units within a cluster may be similar with regard to characteristics of interest, so that including an extra sample unit within the cluster improves the precision of the estimate much less than does including an extra unit outside the cluster. In other words, a sample of \( n \) units made up of a series of \( m_i \) units in \( i \) clusters will be less efficient than a simple sample of \( n \) units dispersed randomly in the population.
The loss of efficiency attributable to cluster sampling can be assessed by computing the intraclass correlation coefficient. It is defined as

$$\delta = \frac{\sigma_c^2 - (\sigma^2/M)}{(M - 1) \cdot (\sigma^2/M)}$$

where $\sigma_c^2 = \text{variance between clusters}$

$\sigma^2 = \text{total variance}$

$M = \text{total number of units within a cluster}.$

If $M$ is large, an approximation to $\delta$ is given by $\sigma_c^2/\sigma^2$—that is, the ratio of the between-cluster variance to the total variance.

The relative efficiency of a simple random sample compared with a cluster sample is given by

$$z = 1 + \delta (\bar{m} - 1)$$

where $\bar{m}$ is the number of units selected within a cluster.

Studies have shown that, under a wide variety of conditions, characteristics such as crop yields and the area of specific crops grown per holding have a value of $\delta$ of the order of at least 0.2–0.3. Indeed, values as high as 0.5 have been observed. The implications for sampling efficiency can be seen from the following table showing values of $z$:

<table>
<thead>
<tr>
<th>$m$</th>
<th>0.2</th>
<th>0.3</th>
<th>0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>1.8</td>
<td>2.2</td>
<td>3.0</td>
</tr>
<tr>
<td>10</td>
<td>2.8</td>
<td>3.7</td>
<td>5.5</td>
</tr>
<tr>
<td>20</td>
<td>4.8</td>
<td>6.7</td>
<td>10.5</td>
</tr>
</tbody>
</table>

With an intraclass correlation of 0.2 and a sample size per cluster of five, the clustered sample would need to be 1.8 times larger to give the same precision as a simple random sample of the same size. When the intraclass correlation is 0.3, with a sample size of twenty per cluster, the cluster sample is 6.7 times less efficient, a staggering order of magnitude.

This relationship has serious implications for crop cutting and, to a lesser extent, for the harvest unit method. The only justification for choosing crop cutting is the potentially high level of accuracy that it can achieve. But if available resources limit the sample design
to a relatively small number of clusters, each containing ten or more respondents, the gain in measurement accuracy from the method will be canceled by the loss in sampling precision. The same argument applies to the harvest unit method. However, the harvest unit method is less demanding than crop cutting, and the enumerator may be expected to cope with a more dispersed sample.

**Sampling Units and the Calculation of Crop Yields**

It must be stressed that in samples where the holding is the unit of study, crop yield for the holding should be calculated by dividing the sum of plot outputs by the sum of plot areas. It is not correct to take the simple arithmetic mean of yields per plot. The calculation of mean crop yields for the set of holdings in the sample, or for subpopulations such as holdings on which fertilizer is used, depends upon the purpose for which the average is required. If the intention is to present the mean achievement of project beneficiaries, the required average can be calculated from the simple arithmetic mean of holding yields. If the intention is to contrast crop areas with different characteristics, such as areas with or without inorganic fertilizer, the yields should be computed by dividing the sum of output by the sum of area.

If crop yield is correlated with plot area, use of the simple arithmetic mean in the wrong situation can lead to a significant bias.

**Reporting Results with Mixed Cropping**

Cropping patterns, particularly in Africa, are often complex. Crops are grown singly, in fixed mixtures with other crops, and in relay mixtures. Land is cropped as often as three times a year, with the same crop appearing more than once. Providing a simple summary of the area and the yield of a particular crop presents considerable difficulties.

There are two basic types of crop mixture: (1) one crop is occupying space within the plot that would otherwise be occupied by another, so that each crop is grown at a lower density than would occur if they were grown separately; (2) one crop is added between the rows of another crop, which has been planted at its normal density. Clearly, crop production is a function not simply of plot
area but also of the relative plant density in each mixture and the detrimental or beneficial effects of the other crops in the mixture. If crop area is presented as a simple sum of all land on which the crop appears, irrespective of mixture, the resulting figure will be misleading unless it is supported by other information. A number of different alternatives have been used in an attempt to overcome this problem—either by standardizing crop areas to a common base or by preparing specially constructed tables.

The proposals for standardization involve converting, by one means or another, the area under a crop mixture into an equivalent area devoted to a single crop. In the simplest example, areas of crop mixtures are divided by the number of crops in the mixture. Thus 1 hectare of maize and beans would count as 0.5 hectare of maize and 0.5 hectare of beans. Another method gives the whole area to each constituent crop. These oversimplistic methods have been widely used. Most project reports of crop areas as well as international crop statistics are presented as if the crops were grown in pure stand, which reflects an attempt to make a complex reality fit simulation. Various refinements have been proposed: standardizing the mixture to the sole crop by seed rate, plant density, and so on. Another method is to give the total area to a so-called main crop and assign varying proportions to other constituents. These methods require a double calculation (weighting the plot area by the mixed-crop characteristic standardized to the sole-crop characteristic, then reapportioning crop areas so that the areas sum to the total) to ensure that the resulting sum of individual crop areas will reflect the cropped area total. The main weakness is not the tortuous calculations involved, but the problems inherent in choosing a standard to act as the denominator. Sole-crop densities, yields, and seed rates vary considerably from year to year and from one region to another. The surveyor could, of course, attempt to report every mixture in detail, but this is not practical. Even where only a handful of sole crops and major mixtures occur, many hundreds of additional mixtures can be found. To present even a portion of these is to give a spurious impression of accuracy and to overwhelm the reader with unnecessary detail. Moreover, huge samples are required to achieve even modest precision.

The most reliable approach, and therefore the recommended approach, is to present at least two levels of detail: first, the overall
land area on which the principal crops are grown, together with crop yields; and second, for each crop a breakdown of the area into certain basic types—for example, maize in pure stand, maize with other cereals, maize with beans and pulses, maize with permanent crops, maize with all other crops. The fact that these cannot be aggregated over crops to give total cropped area (because of double counting) is of no concern since the cropped area is presented separately.

Scale of Inquiry

A recurring theme in both the Casley and Lury handbook and this technical supplement is the idea that small may be better, especially when high standards of accuracy are required for the subsequent analysis. In this context the issue is not the size of samples chosen for M&E surveys, but the scale of operation of the M&E survey unit. A small enumerator force that includes a core of experienced and well-trained staff is easier to manage and is capable of more flexible work programs than larger forces that require more than one level of supervision.

With a smaller team, the M&E officer can be directly involved in the data collection. The burden of analysis is then shifted from reliance on statistical precision from large surveys of uncertain quality to inference which is founded on the field observations and firsthand experience of the M&E team.
APPENDIX: Standardization of Crop Cut Data

In the Nigerian study a series of 100-square-meter squares were randomly located in a set of fields containing either sorghum or yam. The demarcated squares were each divided into a pair of 50-square-meter right-angled triangles. Differences between the estimated yield from each subplot and the actual yield of the field in which it was laid were standardized according to the actual yield. The bias was estimated in terms of the departure of the mean of the distribution of standardized differences from zero. The result was an estimated bias of 14 percent for each crop.

An analysis of the comparative biases between the 50-square-meter triangle and the 100-square-meter square is made difficult by the lack of independence between the triangles and the squares. Because of the method of selection, the mean estimates from the triangles and the squares are the same. A variant of the standardization method was used—standardizing the differences between the subplot estimate and the actual yield by the standard deviation of the estimates rather than the actual yield. The resulting analysis showed mean errors of the order of 8–10 percent, with no significant difference between the 50-square-meter and 100-square-meter subplot sizes—although the 100-square-meter subplot bias was slightly higher for both sorghum and yam.
The most recent World Bank publications are described in the annual spring and fall lists. The latest edition is available free of charge from the Publications Sales Unit, Department B, The World Bank, Washington, D.C. 20433, U.S.A.
This booklet examines the relative merits and limitations of various methods for measuring crop production on smallholder farms in agriculture and rural development projects. It focuses in particular on measurement techniques used in estimating annual crop yields. In their detailed analysis of one of the most popular techniques, crop cutting, the authors draw on heretofore unpublished studies from northern Nigeria.

The authors stress that the choice of measurement technique depends on the type of survey to be undertaken, and they offer specific, practical recommendations that are designed to help monitoring and evaluation staff make the appropriate selections. The booklet is intended as a technical supplement to Dennis J. Casley and Denis A. Lury, Monitoring and Evaluation of Agriculture and Rural Development Projects (Johns Hopkins University Press, 1982), the popular how-to handbook on the design and implementation of monitoring and evaluation systems.