Measuring coffee yields: A comparison of one-time harvest and allometric methods

Vivian Hoffmann, Godfrey Kagezi, Ezra Rwakazooba, Giulia Zane

1. Introduction

Measuring yields accurately is critical for evaluating the impact of interventions that aim to increase agricultural productivity. Self-reported survey data may be subject to social desirability bias, especially in the context of intervention-based studies. Further, farmers participating in training programs may put more effort into measuring their yield because they are interested in learning whether the new practices are effective, which could systematically bias their self-reports, compared to those of farmer not offered training. Therefore, it is often argued that field measurements such as crop cuttings and weighing are required for accurate measurement.

In the case of coffee, the harvest season spans several months. Under natural conditions, coffee fruits (cherries) ripen at different points in time, even within the same tree (Da Matta et al., 2007). The most accurate way to measure coffee yields is thus through multiple harvests of whole selected plants over the course of a fruiting season (e.g. Bote and Vos, 2016).

However, given the high cost of multiple farm visits, a more common approach is to multiply one-time pre-harvest counts or estimates of a sample of cherries per plant by the average weight of the ripe fruits or beans (Soto-Pinto et al., 2000; Cilas and Descroix, 2009; Idol and Youkhana, 2019). This method, referred to as allometric measurement, has the additional advantage of being non-destructive and is expected to be acceptable to most farmers. A potential shortcoming of allometric measurement is that accurately counting cherries requires careful attention to detail. In the context of large surveys in which enumerator effort cannot be directly monitored, enumerators may systematically under- or over-count cherries, leading to bias of unknown direction.
A third yield measurement option that combines these two approaches is to harvest all the ripe and unripe fruit at one time, and sort these into categories and report the per-category weights (Campanha et al., 2004). Like the multiple yield-based method, this approach has the advantage of being less affected by potential enumerator attention constraints. However, this approach has two drawbacks. First, stripping coffee trees of all leaves and cherries is known to increase susceptibility to disease and reduce harvest in subsequent years (Willsion, 1999 as cited by Daniels, 2009). While a one-time harvest can be done more carefully and selectively than full stripping, the yield impact of this practice is not known. Second, the acceptability of picking cherries ahead of the harvest in the study context of western Uganda is not known. Most coffee farmers in this region operate at small scale and many sell their harvest ahead of maturity. Harvesting a subset of trees ahead of the harvest could potentially create conflicts with buyers who have a claim on production of these trees.

In this note, we assess the suitability of allometric measurement for use in the context of large-scale data collection on coffee yields. We compare allometric yield estimates of Robusta coffee grown by smallholder farmers in Uganda with those obtained through a one-time harvest of both ripe and unripe cherries of a specified minimum size at the beginning of the fruiting season.

2. Materials and methods

2.1 Sample selection

The study sample consists of coffee farmers in 22 villages of Mateete subcounty within Ssembabule district, in Western Uganda. Villages were randomly selected from a list of villages identified by a non-governmental organization as appropriate for a training program on coffee agronomy. The NGO had not conducted any intervention in the study villages at the time of the study.

In each selected village, a listing of coffee farmers was conducted with the help of community leaders. Ten farmers were then randomly selected in each village to be included in the study, and informed consent was obtained to conduct and interview and inspection of the coffee farm. In some villages, the number of coffee farmers available to be interviewed was fewer than 10, resulting in a final sample size of 213 farmers.

2.2 Data collection

Data collection was conducted March and April of 2018, prior to the beginning of the main coffee harvest in the region. The following protocol was used to select coffee trees.

First, the enumerator measured the area of the coffee field using a hand-held GPS unit. If the area of the field was above half an acre, coffee trees on four separate 10x10 meter squares would be inspected. For smaller fields, only three squares would be used.

Next, the enumerator walked the longest diagonal connecting two opposite corners of the field, counting his or her steps. Three or four randomly selected stopping points were drawn by the data collection program. Each stopping point constituted the northwest corner of a 10x10 meter sample square. In each square, the number of productive coffee trees was counted for use in estimating density and area-
based yields. Allometric yield measurement was conducted on the trees closest to the northwest and southeast corners of the square.

Consent to harvest the ripe and unripe cherries form two coffee trees, in exchange for monetary compensation (set at higher than the market price of coffee to compensate for any potential negative impact on future yield), was obtained separately from consent to participate in the study. All farmers except one agreed to allow the researchers to strip their trees. For those who agreed, all cherries of at least 11 mm length on the trees closest to the northwest corner in the first two sampling squares were removed for weighing.

2.3 Harvest-based yield measurement

This method involved removing all cherries at least 11 mm in length from the tree and then dividing them into size groups to be weighed separately. Any harvested cherries found to be less than 11 mm in length were discarded prior to weighing, as these were not expected to reach maturity during the coming harvest season. Further, at this stage, cherries may fail to ever reach maturity, and thus counting them would over-estimate yield.

Cherries larger than this size were grouped into three categories based on length: 11 to 12.5 mm, 12.5 to 14 mm, and longer than 14 mm. The share of all cherries in each size group was calculated, as was the share of red cherries in each size group.

2.4 Allometric yield measurement

This methodology consists of counting and then randomly selecting from among the main stems, branches, and clusters of cherries on the tree, and then counting the number of cherries in each selected cluster. We followed Cilas and Descroix (2009) in sampling one stem, four branches on that stem, and two clusters of cherries per branch.

Only productive stems, defined as those with at least one groundnut-sized cherry on it, were counted. These stems were labeled with papers indicating numbers from one up to the number of stems, and a time-stamped photograph was taken. This was done to avoid the risk of enumerators purposively selecting stems that contain fewer cherries. Next, the number of branches on the stem (whether productive or not) were counted. This number was entered into a tablet computer, and four branches were randomly selected by the data collection program. The clusters of cherries on each of the selected branches were then counted, this number entered, and two clusters per branch were selected. Finally, the number of cherries at least 11 mm in length on each of the randomly selected clusters were counted.

3. Analysis

3.1 Estimation of yield per tree
For both methods, the first step in calculating yield was to estimate the number of cherries per tree. For each stripped tree, the weight of cherries harvested in each size category was divided by the weight of a single cherry in that category. This provided an estimate of the total number of cherries on the tree.

From the allometric data, the number of cherries counted per cluster was multiplied by the total clusters counted on sampled branches, the total branches counted on sampled stems, and the total stems counted on the tree.

The estimated yield was then calculated according to the following formula:

$$\text{estimated yield} = \text{total cherries} \cdot \sum_{i=1}^{3} \text{share of red cherries in size category } i \cdot \text{weight of a category } i \text{ cherry}$$

As in the harvest protocol, cherries smaller than a small groundnut were not included in this estimate.

3.2 Comparison of harvest-based versus allometric yield estimates

Given the highly skewed distribution of coffee yields, we examine the geometric mean of the yield estimates (kg of red cherry). Visual inspection results in a yield that is qualitatively lower, but statistically indistinguishable from the harvest-based yield for the first tree (0.118). The second tree harvested was the fourth tree visually inspected, as not all visually inspected trees were harvested. For this tree, the allometric yield estimate is significantly lower than the harvest-based yield for the second tree (p=0.009). We interpret this finding as indicating flagging enumerator attention to detail over time.
Table 1: Yield estimates (ripe cherry / tree) based on harvested weight and visual inspection, by tree

|                  | Tree 1 |         |         | Tree 2 |         |         |
|------------------|--------|---------|---------|--------|---------|---------|
|                  | harvest| visual  | diff    | p-value| harvest | visual  | diff    | p-value |
| Geometric mean   | 0.85   | 0.73    | 0.12    | 0.118  | 0.80    | 0.59    | 0.21    | 0.009   |
| Stand. deviation | 2.31   | 3.43    |         |        | 2.84    | 3.25    |         |        |
| 25th percentile  | 0.28   | 0.23    |         |        | 0.23    | 0.14    |         |        |
| Median           | 0.99   | 0.81    |         |        | 0.81    | 0.51    |         |        |
| 75th percentile  | 2.34   | 2.51    |         |        | 2.63    | 1.61    |         |        |
| 95th percentile  | 6.08   | 9.24    |         |        | 6.67    | 6.3     |         |        |
| 99th percentile  | 11.6   | 19.6    |         |        | 9.6     | 9.9     |         |        |
| Maximum          | 14.1   | 21.7    |         |        | 29.2    | 35.9    |         |        |
| N                | 212    | 213     |         |        | 212     | 213     |         |        |

The distribution of yield (kg ripe cherry / tree) based on the mean of the two trees for which is both measures were taken, is shown in Figure 1. The harvest-based measure of mean yield (GM = 1.162) is higher than the allometric measurement (GM = 0.958) at p=0.018.

Figure 1: Distribution of yield estimates (mean of 2 trees) based on harvest and allometric methods
4. Conclusions

Our harvest-based approach to measuring coffee yield is similar to that of Campanha and colleagues, but we take the additional step of estimating the number of cherries harvested, to allow comparability with allometric results. We find that the harvest-based measure gives higher estimated coffee yield per tree, and that the difference is increasing in the order of trees for which allometric yields are estimated, consistent with an enumerator fatigue effect.

We find that feasibility of the one-time harvest method in the study site of western Uganda was widely acceptable to farmers. Only one of 213 farmers declined to allow researchers to harvest all of the red and green cherries measuring at least 11 mm from two coffee trees, in exchange for compensation. However, before applying approach to yield estimation at scale, the impact of harvesting unripe coffee cherries on future yields should be evaluated.

ABOUT THE AUTHORS

Vivian Hoffmann is Research Fellow in IFPRI’s Markets, Institutions and Trade Division.

Godfrey H. Kagezi is a Senior Research Officer/Plant Entomologist and the Programme Leader, Coffee Cocoa Plant Health Management Programme (CPhM-p) at the National Coffee Research Institute (NaCORI), within the National Agricultural Research Organization (NARO) of Uganda.

Ezra Rwakazooba is a Managing Partner and Data Scientist at Ignosi Research.

Giulia Zane is an Impact Evaluation Consultant at the Inter-American Development Bank.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the contributions of Siddhartha Baral (IFPRI Research Analyst) and Janet Ayebare (IFPRI Research Coordinator) to data analysis and collection of additional data, respectively.
REFERENCES

Bote, A.D., Vos J. 2016. Branch growth dynamics, photosynthesis, yield and bean size distribution in response to fruit load manipulation in coffee trees. *Trees* 30(4):1275-1285.

Campanha, M.M., Santos, R.H.S., De Freitas, G.B., Martinez, H.E.P., Garcia, S.L.R. and Finger, F.L., 2004. Growth and yield of coffee plants in agroforestry and monoculture systems in Minas Gerais, Brazil. Agroforestry systems, 63(1), pp.75-82.

Da Matta, F.M., Ronchi, C.P., Maestri, M., Barros, R.S. 2007. Ecophysiology of coffee growth and production. *Brazilian Journal of Plant Physiology* 19(4):485-510.

Cilas, C. and Descroix, F. 2009. Yield estimation and harvest period. In: Wintgens JN (ed) *Coffee: growing, processing, sustainable production. A guidebook for growers, traders, and researchers*, 2nd edition. Wiley, Weinheim, pp 601–609.

Daniels, N., 2009. *Variations in coffee processing and their impact on quality and consistency* (M.Sc. thesis, Michigan Technological University). 80 pp.

Idol T.W. and Youkhsana AH, 2019. A rapid visual estimation of fruits per lateral to predict coffee yield in Hawaii. *Agroforest Systems* pp. 1-13. [https://doi.org/10.1007/s10457-019-00370-y](https://doi.org/10.1007/s10457-019-00370-y).

Willson, K. C. (1999) *Coffee, Cocoa, and Tea*. CABI Publishing, New York, 300p.

Funding for this work was provided by Benckiser Stiftung Zukunft, and by the CGIAR Research Program on Policies, Institutions and Markets (PIM) led by IFPRI. This publication has not been independently peer reviewed. Any opinions expressed here belong to the author(s) and are not necessarily representative of or endorsed by IFPRI.