

ABSTRACT

Milk production per lactation (MPL) is a key metric of dairy farms. Accurate estimation of MPL requires regular recording, which is laborious and costly. In smallholder systems in the tropics, therefore, generally very few records are available to estimate MPL. Cross-sectional studies collect only one single record per lactation, and even longitudinal studies usually yield only a limited number of records per lactation. Such data recording methods, therefore, are sometimes extended with records recalled by farmers. The accuracy of MPL-estimates based on such limited and imperfect data, however, is unknown. The aim of the present study was to assess the accuracy of MPL-estimates from a single record and a limited number of records per lactation, obtained from smallholder dairy farms in Nakuru County, Kenya. Test-day records from a milk recording scheme for 114 smallholders were used to prepare three datasets with: i) a complete number of test-days (CTD, 5803 records), ii) a limited number of test-days (LTD, 1583 records), and iii) a single test-day (STD, 471 records). In addition, farmers' recall data (i.e. farmers retrieve information from the past, through memory) from a survey of 29 farms with 56 lactations were used to prepare two datasets with: i) a limited number of recall moments per lactation (LRM, 200 records), and ii) a single recall moment per lactation (SRM, 56 records). These five datasets were used to derive MPL-estimates, at individual cow level or at herd level. The latter was done to mimic a situation without individual cow data, but only herd data (i.e. yield and size). MPL-estimates for CTD were set as a benchmark to quantify the accuracies, based on the relative mean absolute error (RMAE) and root mean square error (RMSE), of MPL-estimates for LTD and STD. As a benchmark dataset was absent for recall data, we computed a virtual benchmark to quantify the accuracies of MPL-estimates for LRM and SRM. At cow level, accuracy of MPL-estimates was highest for LTD (RMAE 15%), and lowest for SRM (RMAE 28%), while accuracies for STD and LRM were intermediate (RMAEs ~ 20%). At herd level, accuracy was higher for STD (RMAE 13%) than for SRM (RMAE 25%). We also showed that to detect a difference of, for example, 100 kg in MPL we need 3002 cows for CTD, and between 3620 and 5003 cows when using alternative data collection methods. Hence, depending on the study objective, alternative data recording methods provide labor-saving and cost-effective ways to estimate MPL in data-scarce smallholder dairy systems.