Data Descriptor

Title

The Rural Household Multiple Indicator Survey (RHoMIS) data of 13,310 farm households in 21 countries.

Authors

Mark van Wijk[1*#], James Hammond[1#], Leo Gorman[2#], Sam Adams[1], Augustine Ayantunde[1], David Baines[3], Adrian Bolliger[4,5], Caroline Bosire[1], Pietro Carpena[3], Sabrina Chesterman[6], Amon Chinyophiro[7], Happy Daudi[8], Paul Dontsop[9], Sabine Douxchamps[4], Willy Desire Emera[1], Simon Fraval[1,10], Steven Fonte[11], Lyda Hok[12], Henry Kiara[1], Esther Kihoro[1], Luke Korir[1], Christine Lamanna[6], Chau TM Long[13], Godfrey Manyawu[1], Zia Mehrabi[14], Dejene K Mengistu[15], Leida Mercado[16], Katherin Meza[11,17], Vesalio Mora[15], Jacob Mutemi[18], Mary Ng'endo[6], Paulin Njingulula[9], Chris Okafor[9], Tim Pagella[19], Phonepaseuth Phengsavanh[20], James Rao[1], Randall Ritzema[1,21], Todd S. Rosenstock[6], Tom Skirrow[3], Jonathan Steinke[15], Clare Stirling[22], Jose Gabriel Suchini[16], Nils Teufel[1], Peter Thorne[1], Steven Vanek[11], Jacob van Etten[15], Bernard Vanlauwe[9], Jannike Wichern[23], Viviane Yameogo[1]

Affiliations

1. International Livestock Research Institute (ILRI), PO Box 30709, Nairobi 00100, Kenya
2. Alan Turing Institute, British Library, 96 Euston Road, London, England, NW1 2DB
3. TreeAID, Brunswick Court, Brunswick Square, Bristol, BS2 8PE, UK
4. International Center for Tropical Agriculture (CIAT), Asia Regional Office, Hanoi, Vietnam
5. Hanns R. Neumann Stiftung, Kota Liwa 34812 Lampung Barat, Indonesia
6. World Agroforestry (ICRAF), PO Box 30677-00100, UN Avenue, Nairobi, Kenya
7. NASFAM, African Unity Avenue, Lilongwe, Malawi.
8. Tanzania Agricultural Research Institute – Naliendele, Mtwara,Tanzania
9. International Institute of Tropical Agriculture (IITA), P.O. Box 30772-00100, Nairobi, Kenya
10. Animal Production Systems Group, Wageningen University & Research, P.O. Box 338, Wageningen 6700 AH, the Netherlands
11. Department of Soil and Crop Sciences, Colorado State University, Fort Collins, CO, 80523, USA
12. Faculty of Agronomy, Royal University of Agriculture, Phnom Penh 12401, Cambodia
13. Long Western Highlands Agriculture and Forestry Science Institute (WASI), Buon Ma Thuot, Vietnam
14. The UBC School of Public Policy and Global Affairs, University of British Columbia, Vancouver, British Columbia V6T 1Z2, Canada
15. Bioversity International, Rome, Italy
16. Centro Agronómico Tropical de Investigación y Enseñanza (CATIE), Turrialba, Costa Rica
17. Grupo Yanapai, Calle Tripoli 365, Miraflores, Lima, Peru
18. Lutheran World Relief, Nairobi, Kenya
19. School of the Environment, Natural Resources and Geography, Bangor University, UK
20. National Agriculture and Forestry Research Institute (NAFRI), Vientiane, Lao PDR
21. Olivet Nazarene University, One University Avenue, Bourbonnais, IL60914, USA
22. International Maize and Wheat Improvement Centre (CIMMYT), Sustainable Intensification Program, Mexico DF, Mexico
23. Plant Production Systems group, Wageningen University, P.O. Box 338, 6700 AH Wageningen, The Netherlands

* Corresponding author(s): Mark van Wijk (m.vanwijk@cgiar.org)
# Contributed equally to this study

Abstract

The Rural Household Multiple Indicator Survey (RHoMIS) is a standardized farm household survey approach which collects information on 753 variables covering household demographics, farm area, crops grown and their production, livestock holdings and their production, agricultural product use and variables underlying standard socio-economic and food security indicators like the Poverty Probability Index, the Household Food Insecurity Access Scale and dietary diversity. These variables are used to quantify more than 40 different aggregate indicators on farm household characteristics, welfare, productivity and economic performance. Between 2015 and the beginning of 2018, the survey instrument has been applied in 21 countries in Central America, sub-Saharan Africa and Asia. The data presented here cover the raw data, the indicator calculation code and the resulting indicator values, and can be used to quantify on- and off-farm pathways to food security, diverse diets and reduced poverty of rural smallholder farm households.

Background & Summary

Agriculture is the most important livelihood option for most rural households in low and middle-income countries[1]. Smallholder farm households in these locations produce food not only for themselves, but, in many countries, produce the majority of the national, or even regional food supply[2,1]. Smallholder farm households are also highly diverse[3], varying in size and amount of livestock present, crops grown, and farm management strategies. The importance of understanding the diversity and dynamics of rural households is increasingly crucial, given the diverse effects of global changes in climate, population growth, urbanization, and food demand[4,5].

Achieving the sustainable development goals (in particular on no poverty and zero hunger, but others too) requires more intensified sustainable food production and development of rural economies. Targeted investment to make progress in agricultural development requires understanding the links between farming and livelihood practices, and their effects on farm performance and household welfare. Reliable indicators at farm-household level of both farm performance and household welfare are therefore needed to better understand and model these linkages, and to inform the design and implementation of interventions by governments, donors, and international agencies, across a wide range of differing geographies and socio-economic dimensions[6].

The lack of standardization of agricultural household surveys, especially in international 'agriculture for development' research, has resulted in a proliferation of survey tools and development indicators. The consequence is incoherent data. This limits our ability to compare outcomes across studies and to draw general conclusions on intervention effectiveness and trade-offs between outcomes, which may be shaped by household structure, farm management and the wider social-environmental context[3,5]. In contrast, RHoMIS (Rural Household Multiple Indicator Survey; www.rhomis.org) is a standardized household survey approach designed to rapidly characterise a series of key indicators across the spectrum of agricultural production and off farm activities, alongside market integration, nutrition, food security, poverty and greenhouse gas (GHG) emissions[7] (Figure 1). It includes a modular survey tool which takes 40-60 minutes to administer per household, a digital platform to store and aggregate incoming data as well analysis code to quantify indicators and visualize results. The tool has been systematically designed to enable the quantification of interactions between different indicators of agricultural systems, productivity, and human welfare at the farm and household level and has been widely adopted by research organisations and development partners[8].

A harmonized dataset has been developed from all the applications of RHoMIS that took place during the years 2015, 2016, 2017 and the first three months of 2018, resulting in a data collected from than 13,310 farm households across 21 low and middle income countries. The database consists of the raw data (the 753 variables mentioned above; see subdirectory 'data to share\rhomis_full_data.csv') and 41 indicators calculated based on the information provided by these variables (see subdirectory 'data to share\rhomis_indicators.csv'). The raw data and indicators have already been used for a wide range of studies at site level[8,9,10,11,12,13], for regional analyses [14,15], and for continental analysis[16]. Different aspects of smallholder households have been analysed, including gender equity[14,17], dietary diversity[15,16], nutritional gaps[16], poverty and GHG emissions in relation to production intensification[7], subsistence- versus market-orientated strategies[16], and on-farm vs. off-farm activities[16]. RHoMIS is an on-going initiative, and we welcome interested parties to the community of practice (see www.rhomis.org for up-to-date information and downloadable survey questionnaires). Records continue to be submitted to the central data repository: in the latter part of 2018 more than 10,000 households were additionally interviewed, and their information added to the database. Further releases will be made public in the near future.

Figure 1. Key indicator groups generated by all RHoMIS applications

Methods

Basic characteristics and geographic coverage

The countries in which survey data included in the current database were collected, are summarized in Table 1, together with key information for each survey application: the research or development project in which the survey was applied, the lead organisation implementing the survey and the number of households surveyed. The geographical locations of the surveyed households is shown in Figure 2, showing the widespread application of the survey. The public version dataset includes detailed documentation and metadata ('MetaData RHoMIS survey applications.docx'), describing the aims of each project in which RHoMIS was applied, the sampling strategy used, and additional data collected outside the core set of RHoMIS variables (i.e., Figure 1). These additional, project specific data that are outside the core RHoMIS variables are available on request from the corresponding author of this manuscript, always in consultation with the representative of the organisation which executed the RHoMIS application in the field.

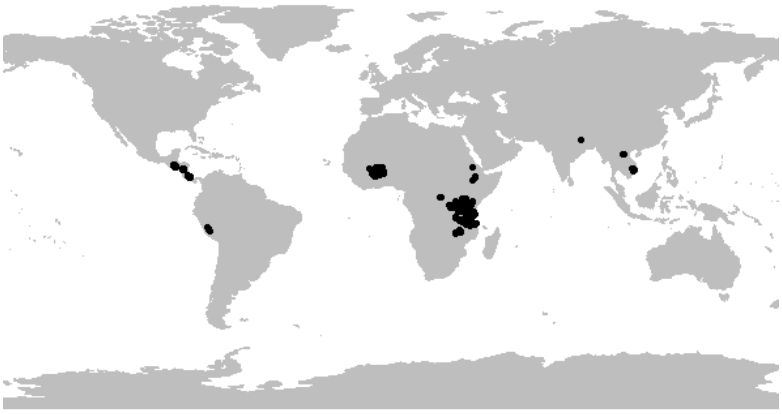


Figure 2. Geographical location of the observations included in the RHoMIS dataset

Table 1: RHoMIS applications included in this data set

Country (site/region)	Lead Institute	Year	Households surveyed
Tanzania	International Livestock Research Institute (ILRI)	2015	150
Trifinio (Guatemala, El Salvador, Honduras)	Bioversity International	2015	300
Mali	International Livestock Research Institute (ILRI)	2015	200
Burkina Faso	International Livestock Research Institute (ILRI)	2015	200
Malawi	International Maize and Wheat Improvement Centre (CIMMYT)	2015	160
Kenya (Central-West)	International Livestock Research Institute (ILRI)	2016	160
Kenya (South)	International Livestock Research Institute (ILRI)	2016	160
Burkina Faso	International Livestock Research Institute (ILRI)	2016	400
India	International Livestock Research Institute (ILRI)	2016	160
Cambodia	International Center for Tropical Agriculture (CIAT)	2016	631
Vietnam	International Center for Tropical Agriculture (CIAT)	2016	310
Laos	International Center for Tropical Agriculture (CIAT)	2016	365
Tanzania	International Livestock Research Institute (ILRI)	2017	996
Tanzania	Bioversity International	2017	528
Ethiopia	Bioversity International	2017	249
Kenya	Bioversity International	2017	316
Democratic Republic of Congo	World Agroforestry (ICRAF)	2017	413
Kenya	World Agroforestry (ICRAF)	2017	385
Zambia	World Agroforestry (ICRAF)	2017	600
Democratic Republic of Congo	International Institute of Tropical Agriculture (IITA) - International Livestock Research Institute (ILRI)	2017	438
Ethiopia	TreeAID	2017	302
Ghana	TreeAID	2017	223
Mali	TreeAID	2017	363
Uganda	Wageningen University	2017	130
Kenya	International Livestock Research Institute (ILRI)	2017	169
Costa Rica	Bioversity International	2017	204
Burundi	International Livestock Research Institute (ILRI) - International Institute of Tropical Agriculture	2018	333
Burkina Faso	TreeAID	2018	1030
Tanzania	World Agroforestry	2018	841
Peru	Grupo Yanapai - Colorado State University	2018	174
Nicaragua	Bioversity International	2018	145
Burkina Faso	TreeAID	2018	1270
Ethiopia	TreeAID	2018	405
Ghana	TreeAID	2018	391

The total number of households in the data set is 13,310. The data are all based on single cross-sectional survey applications. In most applications, households were chosen randomly within the sites in which the different implementing projects were working (see the metadata document for more information on the different sampling strategies applied). Three country-wide applications are available in the current database: two led by the iNGO TreeAID in Burkina Faso, where households were selected across the major agro-ecological zones of the country, and one led by the International Livestock Research Institute in Tanzania that sampled cattle owning households, randomly selected from animal health service records. As is clear from Table 1, the sample sizes differ per application. It is also of key importance to refer to the meta data information file to assess the representativeness of each of the RHoMIS applications. In previous analyses that aimed for statistical inferences valid for farm type, households were grouped by their farm characteristics[16], or if they aimed for spatial representativeness, observations were spatially clustered and resampled or weighted by the local population (e.g. village) densities[13]. The varied sampling procedures followed by the individual applications make it essential to weight and/or re-sample households in any analysis making use of this combined RHoMIS dataset for valid statistical inferences.

The Questionnaire

The RHoMIS questionnaire is a set of carefully, expertly designed modules that are administered digitally using the ODK software platform (https://opendatakit.org/)[7]. The survey is designed to be both flexible enough to suit local contexts and sufficiently standardised to permit rapid deployment, analysis and comparison between multiple sites, without the need for costly post-survey harmonization. The data package that is made available consists of two parts: the dataset itself (containing the raw data and the indicator results) and secondly, the series of documents and analysis code files underpinning the raw data collection effort and the subsequent indicator quantification. We have made the survey available in easily readable pdf format ('RHoMIS for printing_v1.3.pdf'). The questionnaires and their variable names have been linked to a supporting set of data extraction and analysis tools written in R (https://cran.r-project.org/), also included in the package. The majority of questions in the survey are used for the estimation of a series of pre-defined indicators that include:

- The Household Food Insecurity of Access Scale[18] for measuring the frequency and severity of hunger (this indicator in more recent, 2019 and beyond, applications of RHoMIS is replaced by the FIES – Food Insecurity Experience Scale indicator; http://www.fao.org/in-action/voices-of-the-hungry/background/en/).
- The Household Dietary Diversity Score[19], providing an indicator of household dietary adequacy; this indicator was adapted to cover both the bad and good seasons from a food security perspective
- The Probability of Poverty Index[20,21], an asset-based scoring system to estimate the likelihood that a household is in poverty.
- The Potential Food Availability indicator for quantifying the ability of a household to feed itself through both on-farm and off-farm activities[3,4].

These indicators are combined with a comprehensive inventory of agricultural crops and livestock production characteristics, including yields, the use of products (consumed, sold, etc.), product sales prices and input types, and an assessment of off-farm incomes. The data captured in the RHoMIS tool place the farm household along a continuum of household and farm characteristics, performance indicators and welfare indicators (see Figure 1), enabling in-depth analyses of individual indicators, but also integrative analyses of how indicators co-vary and how on-farm and off-farm livelihood strategies correlate to food security, poverty and dietary diversity.

The questionnaire is organised into seven sections and respondents are asked to provide information on the previous 12 months' farming and non-farming activities (Table 2). A pdf and an excel ODK definition file explaining each variable included in the survey, are also supplied with the data.

Table 2. General overview of the information (captured by the 753 variables shared in the dataset) captured, organised by indicator group (see Figure 1) in RHOMIS

Indicator Group	Questions
Meta data	Informed consent; location of the survey (village, district, region, country); project beneficiary; local monetary unit; GPS coordinates of where the interview took place
Household demographics	Respondent sex; whether respondent is household head; age of respondent; type of household (single female; single male; couple; couple - spouse living away); age and sex of household members
Land use/tenure	Total land area cultivated; amount of land owned; access to common land; land rented out/in
Crop production and use	Crops/ vegetables/trees grown, area grown; production per year/season; use of products (sold, consumed, stored, processed; price of crop products; use of crop residues
Agricultural inputs and management	Use of mineral fertilizer, manure, pesticides, irrigation: type, area and/or amount used; intercropping; agroforestry; medicine use for livestock; fodder purchased for livestock
Livestock	Livestock species kept on farm; production per year/season; use of products (sold, consumed, stored); price of livestock products; fodder used
Wild foods	Relative importance of wild foods for food security; months of the year collected; most important wild food items
Food security and hunger	9 standard questions of Hunger and Food Insecurity Access Scale (HFIAS); number of months with hunger; identification of worst and best month in terms of food security
Household dietary diversity	9 food groups of the Minimum Dietary Diversity Score for Women in the worst, best and last month; source of each food group (farm-based, purchased, gathered or a combination of sources)
Off-farm income	Relative importance of off-farm income in overall income; use of income (reinvestment in farm, household (non-food) expenses, food purchase)
Credit, aid and debts	Household access to credit; dependency on/receipts of aid in the last 12 months, from institutions or neighbours; household debt
Poverty Probability Index (PPI)	10 standard questions, country-specific (see https://www.povertyindex.org/)
Gendered control of produce and income	Per produce and income item: who takes sales' decisions and who controls the benefits (consumption or cash)

Data processing and indicator calculations

A standard set of scripts have been created in R, an open-source data management and analysis system, for processing the raw data provided in this dataset, and for the calculation of the various indicators. An overview of the data processing and indicator calculations is given in Figure 3. The starting point of the procedure is the 'raw collected' data. We have limited the data cleaning step to only correcting obviously unrealistic values, such as when it was clear that the recorded value in combination with the reported unit generated an impossible indicator value (say 1.5 *kg* of

maize yield from a 1 ha field; here it is clear that the unit should be *tons* of maize; or similarly if a production of 3,000 *tons* from a 1 ha field is reported, it is clear that the unit should be *kg*). However, the most important step in the cleaning process is unifying crop names, livestock species names and the crop and livestock product names (step 2 & 3 in Figure 3). This cleaned data is merged into a single raw data csv-file, supplied in the data package (step 6 in Figure 3).

Conversion factors use in the indicator calculations, for example area units, monetary units and energy contents, are based on literature resources (step 7 in Figure 3). Within the calculations of some indicators, e.g. income, value of farm production and crop and livestock productivity, intermediate output variables are also generated: farm gate prices of farm products and crop and livestock production values per individual crop and livestock species (step10 in Figure 3). These are included in the data package. The final product of the indicator calculation is the merged RHoMIS indicator results file, step 13 in Figure 3. The annotated R code, for performing and documenting all the processing and calculation steps illustrated in Figure 3, is also supplied. In addition, the individual indicator calculations are explained in detail in the 'Explanations_of_Calculations_and_Outputs.xlsx' file.

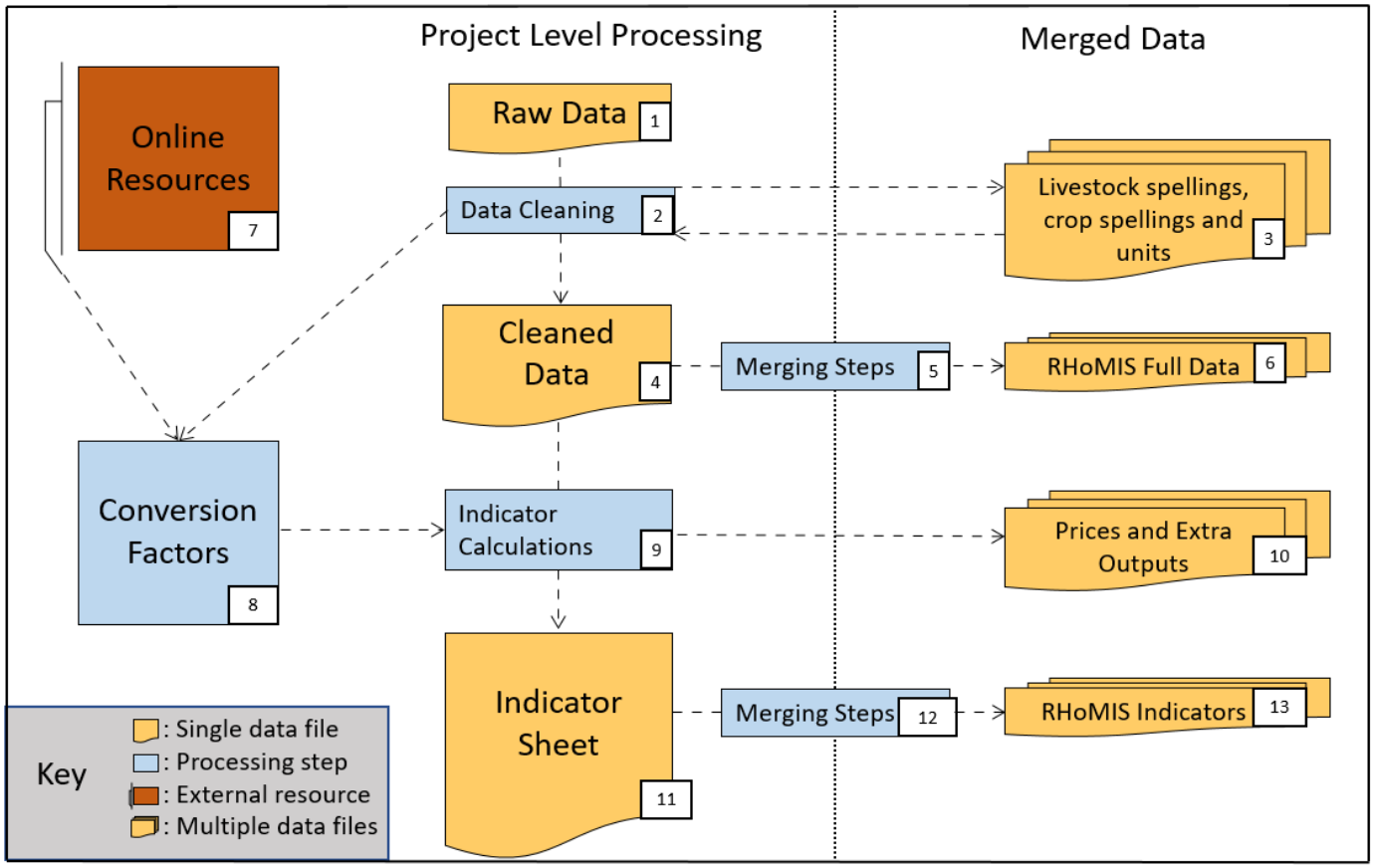


Figure 3. Overview of the data processing and indicator calculation steps.

Data Records

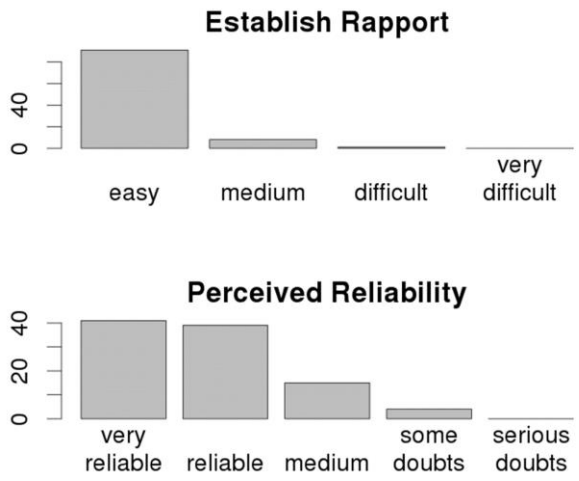
All 753 variables of the survey data are described in the file 'Raw Data code book.xlsx'. The variables included in indicator results file are described in 'RHoMIS Data Processing Doc.docx'. We have also included the RHoMIS survey in ODK definition and pdf format to further facilitate the interpretation of the variable names. The RHoMIS survey is continuously updated; the newest version can be found at www.rhomis.org.

Technical Validation

Each RHoMIS survey application included in the data set has undergone a series of standard data quality evaluation steps to assess the overall quality of the data collected. This did not lead to removal of individual observations, we prefer that the user of the data can make their owned informed decisions on this topic. The validation information

supplied here is only advice on how the data user might be able to check the quality of the data they want to use and to be able to follow systematic approach if he or she does decide to remove data.
In step 1 we check the results of several questions in the survey that are addressed to the enumerators themselves. These questions are 'In your opinion, how easily did you establish rapport with the respondent?' (with possible answers: 'Easily', 'OK', 'Difficult' and 'Very difficult') and 'How reliable do you think these answers are? Consider the accuracy and willingness to answer.' (with possible answers: 'Very reliable', 'Reliable', 'OK', 'Occasional doubts' and 'Regular or serious doubts'). This information in combination with the total survey duration (based on start and end time of survey implementation) provides essential information on the overall reliability of the information recorded. For example, if the survey duration is extremely long (e.g. beyond 2h) it is also likely that data quality will be questionable. Example results of these quality indicators are given in Figure 4, showing the typical distribution of enumerator observed reliability and survey length of survey. This may serve as a useful norm for the initial quality assessment of new RHoMIS applications. The information collected in individual survey applications where the reliability answers show low scores, or where many survey records show abnormally short or long durations, can be more thoroughly investigated or even rejected.

A



B

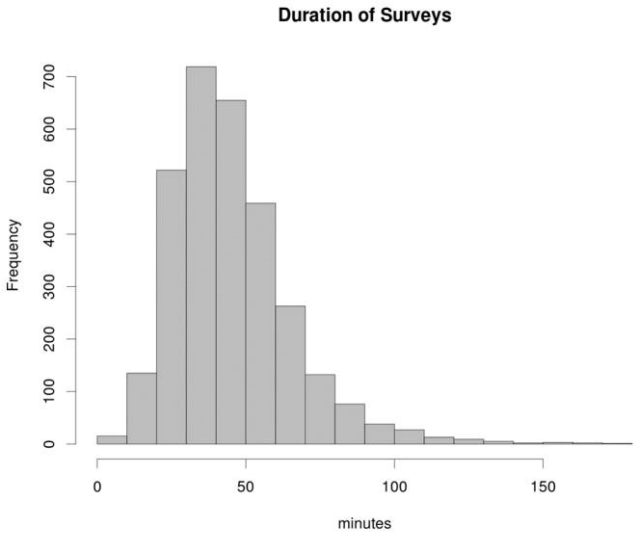
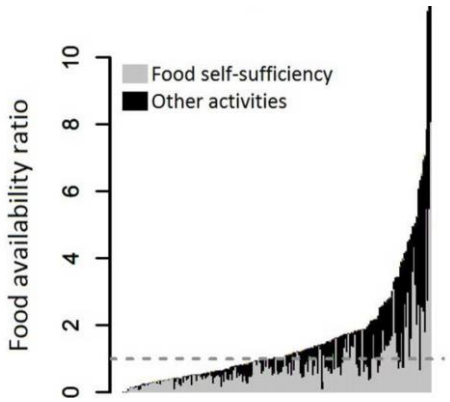


Figure 4. Overview of the two key aspects considered in the initial assessment of overall survey data quality in step 1 (for explanation see text): the enumerator evaluation of data quality (A) and survey duration (B)

In step 2 we follow the food availability/self-sufficiency approach[3,22] to evaluate the overall reliability of the data on production and consumption of farm produce. Calculated livestock productivity and crop yield values are compared to realistic ranges normally encountered in smallholder systems and in the agro-ecosystem of interest [22]. Because the value ranges subsequent to these checks are still large, we also scrutinise so-called composite indicators that combine information from a number of survey variables. The two composite indicators we examine for quality assessment are food self-sufficiency and potential food availability[3]. The food availability indicator represents the total potentially available daily food energy per household member (in adult male equivalent), as provided by the reported consumption of farm products, and by income from the sale of farm products and off-farm activities, whereby all income is converted to a calorific value based on the cost of a local staple crop. Results of these calculations can be used to assess the data quality of information on crop and livestock production, sales and consumption as well as off-farm income. Two problems with this composite indicator are commonly encountered. First, a considerable number of household records at the lower end of the food availability scale appear to suggest an underestimation of calorie availability, suggesting an extreme level of starvation. This may be a true representation of some households, but it can also be an indication of missing information on income or food consumption. Second, a substantial number of households can also show a substantial over-estimation of consumption of crop and livestock products – indicating possible problems with yield, consumption or household size estimates. The lower bound threshold for food availability was set at 1,250 kilocalories (kcal) per male adult equivalent per day, which is below the basal metabolic rate for adult males (approximately 1,590 kcal for a 60 kg male) and adolescents (1,360 kcal for a 40 kg adolescent male). The upper bound for food self-sufficiency were set at: a) 3,500 kcal per male adult equivalent per day,

representing the average intake of developed nations[23]; and b) 5,000 kcal, which is double the approximate requirement for an adult male. The results of these analyses (e.g. Figure 5) are used to examine the overall reliability of the survey application and can be used to identify individual survey applications where the data do not appear reliable from a composite perspective. Typically, between 10 and 25% of the records within the total survey sample of an application site may show questionable values of these indicators. These performance values are representative for the uncertainty encountered in recall based cross-sectional farm household surveys and actually better than the performance of two other widely used sources of farm household information (Figure 5). All sites included in the databases published in this article fall within this performance range. Further trust in these findings can be developed by triangulating these results with other indicators of food security included in the RHoMIS surveys, for example HFIAS, dietary diversity and the number of months with hunger (see also step 3).



	Food availability less than 1250 kcal	FSS above OECD average & less than double	FSS double 2500 kcal	Within bounds
ImpactLite	0.39	0.05	0.04	0.52
RHoMIS	0.15	0.06	0.04	0.75
LSMS-ISA Uganda†	0.12	0.07	0.14	0.66
LSMS-ISA Tanzania†	0.20	0.06	0.07	0.67
LSMS-ISA Ethiopia†	0.40	0.02	0.07	0.51

* Impact lite and RHoMIS include sites from Kenya and Tanzania

†Based on 2010/11 survey round

Figure 5. Step 2 of the data quality evaluation: the identification of systematic under- or over-estimation of production and consumption values based on food self-sufficiency and potential food availability, presented as the relative number of observations. RHOMIS results are compared with results obtained by other survey instruments (ImpactLite and the Living Standards Measurement Survey – Integrated Survey on Agriculture (LSMS-ISA)) in the same countries. Reported in detail in [22]

In step 3 we cross-check relationships between different food security indicators. The information collected by RHoMIS underpins 4 different indicators of food security (HFIAS, HDDS, number of months with hunger and potential food availability). Even though these indicators capture different aspects of food security, we do expect strong correlation between them. Figure 6 provides examples of the typical relationships encountered between these variables. A strong deviation of the relationships found in a new RHoMIS application from these typical relations, may suggest a deeper investigation as to why this is happening, and might indicate a questionable data quality in one or more of these indicators.

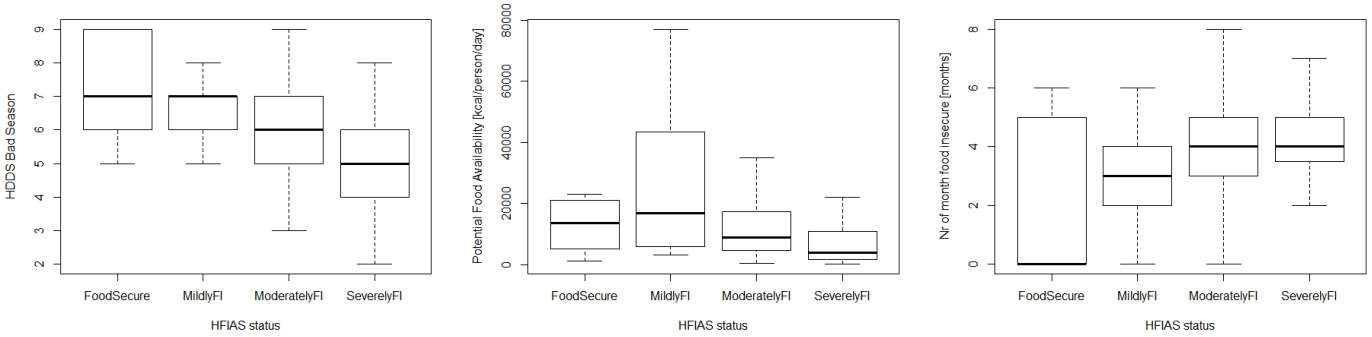


Figure 6. Relationships encountered between the household dietary diversity score in the bad season (HDDS Bad Season), potential food availability and the number of months with food insecurity, and the Hunger and Food Insecurity Access Scale (HFIAS) status, based on the 2016 Kenya Wote dataset

Usage Notes

This data set can be used to investigate the characteristics of agricultural systems in low and middle-income countries at one point in time. On- and off-farm strategies can be analysed in relation to a series of welfare indicators representing food security, poverty and gender equity. This can be done both at individual household level, to disentangle the livelihood strategies of specific household types, but also at population level to determine more generic patterns and investigate equity aspects. We emphasize that the results of such analysis should be interpreted within the context of the data set, considering the potential biases and limitations described in this paper. It is especially important that users note the varied sampling procedures followed in the different RHoMIS applications that it essential to weight and/or re-sample the households in any analysis before making any across-site statistical inferences. See the Methods section for more details.

Acknowledgements

We are grateful to the research teams involved in designing and implementing all surveys made available in this study. Without their rigor, openness and thorough documentation this study would not have been feasible. This study was made possible by the CGIAR Research Program on Livestock and its donors and through the support of the American People provided to the Feed the Future Innovation Lab for Sustainable Intensification through the United States Agency for International Development (USAID). We also thank all donors who made the survey implementations possible through their financial support to the respective projects described in the meta data information file.

Author contributions

JH and MvW designed the RHoMIS tool and coordinated the overall work. JvE, NT, TP and TR gave conceptual input to the RHoMIS tool development. LG prepared the final data, indicator calculations and R scripts. SA coordinated and supported several of the RHoMIS applications presented in the dataset. AA, DB, AB, CB, PC, SC, AC, HD, PD, SD, WE, SF, SF, LH, HK, EK, LK, CL, CTML, GM, KM, DM, LM, KM, VM, JM, MN, PN, CO, PP, JR, RR, TR, TS, JS, CS, JS, NT, PT, SV, JvE, BV, JW, VY have led the RHOMIS applications that form the RHoMIS dataset made available. MvW drafted the manuscript, with critical input from ZM, JH and SF. All co-authors contributed to the finalisation of the manuscript.

Competing interests

The authors declare no competing interests.

Figure Legends

Figure 1. Key indicator groups generated by all RHoMIS applications

Figure 2: Geographical location of the observations included in the RHoMIS dataset

Figure 3. Overview of the data processing and indicator calculation steps.

Figure 4. Overview of the two key aspects considered in the initial assessment of overall survey data quality in step 1 (for explanation see text): the enumerator evaluation of data quality (A) and survey duration (B)

Figure 5. Step 2 of the data quality evaluation: the identification of systematic under- or over-estimation of production and consumption values based on food self-sufficiency and potential food availability, presented as the relative number of observations. RHOMIS results are compared with results obtained by other survey instruments (ImpactLite and the Living Standards Measurement Survey – Integrated Survey on Agriculture (LSMS-ISA)) in the same countries. Reported in detail in [22]

Figure 6. Relationships encountered between the household dietary diversity score in the bad season (HDDS Bad Season), potential food availability and the number of months with food insecurity, and the Hunger and Food Insecurity Access Scale (HFIAS) status, based on the 2016 Kenya Wote dataset.

References

1. Herrero, M., Thornton, P. K., Power, B., Bogard, J. R., Remans, R., Fritz, S., Gerber, J. S., Nelson, G., See, L., Waha, K., Watson, R. A., West, P. C., Samberg, L. H., van de Steeg, J., Stephenson, E., van Wijk, M., & Havlík, P.

(2017). Farming and the geography of nutrient production for human use: a transdisciplinary analysis. The Lancet Planetary Health, 1, e33–e42. doi:10.1016/S2542-5196(17)30007-4.
2. Samberg L H, Gerber J S, Ramankutty N, Herrero M and West P C 2016 Subnational distribution of average farm size and smallholder contributions to global food production Environ. Res. Lett. 11 124010.
3. Frelat, R., Lopez-Ridaura, S., Giller, K. E., Herrero, M., Douxchamps, S., Djurfeldt, A., Erenstein, O., Henderson, B., Kassie, M., Paul, B., Rigolot, C., Ritzema, R. S., Rodriguez, D., van Asten, P. and van Wijk, M. T. (2016) 'Drivers of household food availability in sub-Saharan Africa based on big data from small farms', Proceedings of the National Academy of Sciences of the United States of America, 113(2), 458-463.
4. Ritzema RS, Frelat R, Douxchamps S, Silvestri S, Rufino MC, Herrero M, Giller KE, López-Ridaura S, Teufel N, Paul BK, van Wijk MT. 2017. A simple food availability analysis across smallholder farming systems from East and West Africa: Is production intensification likely to make farm households food-adequate? Food Security 9 (1), 115-131.
5. Waha K, van Wijk MT, Fritz S, See L, Thornton PK, Wichern J, Herrero M. 2018. Agricultural diversification as an important strategy for achieving food security in Africa. Global Change Biology 24(8):3390-3400.
6. Coe R, Sinclair F, Barrios E. 2014. Scaling up agroforestry requires research 'in' rather than 'for' development. Current Opinion in Environmental Sustainability 6, 73-77.
7. Hammond J, Fraval S, van Etten J, Suchini JG, Mercado L, Pagella T, Frelat R, Lannerstad M, Douxchamps S, Teufel N, Valbuena D, van Wijk MT. 2017. The Rural Household Multi-Indicator Survey (RHoMIS) for rapid characterisation of households to inform climate smart agriculture interventions: Description and applications in East Africa and Central America. Agricultural Systems 151, 225-233.
8. Hammond J, van Wijk MT, Pagella T, Carpena P, Skirrow T, Dauncey V. 2019. Shea butter: a pro-poor, pro-female route to increased income. In: Rosenstock T, Nowak, A, Girvetz E (Eds.). The Climate-Smart Agriculture Papers: Investigating the Business of a Productive, Resilient and Low Emission Future. Springer International Publishing
9. Fraval S, Hammond J, Lannerstad M, Oosting SJ, Sayula G, Teufel N, Silvestri S, Poole EJ, Herrero M, van Wijk MT. Livelihoods and food security in an urban linked, high potential region of Tanzania: Changes over a three year period. Agricultural Systems 160: 87-95.
10. Steinke J, Mgimiloko G, Graef F, Hammond J, van Wijk MT, van Etten J. Prioritising household-specific options for agricultural development through the Positive Deviance approach. PlosONE 14(2): e0212926.
11. Steinke J, Achieng JO, Hammond J, Kebede SS, Mengistu DK, Mgimiloko MG, Mohammed JN, Musyoka J, Sieber S, van de Gevel S, van Wijk M, van Etten J. 2019. Household-specific targeting of agricultural advice via mobile phones: Feasibility of a minimum data approach for smallholder context. Computers and Electronics in Agriculture 162, 991 – 1000.
12. Bosire C, Rao J, Hammond J, Lukuyu B, van Wijk MT. Adaptation opportunities for smallholder dairy farmers facing resource scarcity: integrated livestock, water and land management. Agriculture, Ecosystems and Environment 284, 106592.
13. Beveridge L, Whitfield S, van Etten J, van Wijk MT, Hammond, J, Challinor A. Constructing and deconstructing 'food security' across scales: Household indicators and lived experiences in the dry corridor of Central America. Frontiers in Sustainable Food Systems 3, 65.
14. Ritzema RS, Douxchamps S, Fraval S, Bolliger A, Hok L, Phengsavanh P, Long CT, Hammond J, van Wijk MT. Drivers of dietary diversity in transitioning agricultural systems of the Greater Mekong Sub region: Findings from a rapid household assessment in Cambodia, Laos, and Vietnam. Agricultural Systems 176, 102657.
15. Tavenner K, Van Wijk M, Fraval S, Hammond J, Baltenweck I, Teufel N, Kihoro E, deHaan N, van Etten J, Steinke J, Baines D, Carpena P, Skirrow T, Rosenstock T, Lamanna C, Ng'endo M, Chesterman S, Namoi N, Manda L. 2019. Intensifying Inequality? Gendered Trends in Commercializing and Diversifying Smallholder Farming systems in East Africa. Frontiers in Sustainable Food Systems 3, article 10. https://doi.org/10.3389/fsufs.2019.00010.
16. Fraval S, Hammond J, Ayantunde A, de Boer IJM, Bogard J, Bosire C, Chesterman S, Chilingulo K, Chinyophiro A, Dontsop P, van Etten J, Herrero M, Kihoro E, Lamanna C, Lannerstad M, Ng'endo M, Njingulula P, Okafor C, Oosting SJ, Rao J, Rosenstock T, Steinke J, Stirling CM, Teufel N, Vanlauwe B, Waha K, Wichern J, Yameogo V, van Wijk MT. Nutritional gaps of rural households in east and west Africa: prevalence and determinants based on rapid indicators. Frontiers in Sustainable Food Systems, accepted.

17. van Wijk, M., Hammond, J., Fraval, S., Ritzema, R. S., Bolliger, A. M., and Long, C. (2016). "Including gender equity in a survey tool for rural households," in A Different Kettle of Fish. Gender Integration in Livestock and Fish Research, ed R. Pyburn (Amsterdam: LM Publishers), 9–17.
18. Coates, J., Swindale, A. & Bilinsky, P., 2007. *Household Food Insecurity Access Scale (HFIAS) for measurement of food access: indicator guide*, Washington, DC.
19. Swindale, A. & Bilinsky, P., 2006. Household Dietary Diversity Score (HDDS) for measurement of household food access: Indicator guide, Washington, DC. Available at: ftp://190.5.101.50/DISCO 2/SRV-SQL/Nueva carpeta/doc lili bakup6654444/LILY/Monitoreo y Evaluacion/FANTA/fanta HDDS_Mar05 DIVERSIFICACION.doc.
20. Desiere, S., Vellema, W. & D'Haese, M., 2015. A validity assessment of the Progress out of Poverty Index (PPI)TM. Evaluation and Program Planning, 49, pp.10–18. Available at: http://www.sciencedirect.com/science/article/pii/S0149718914001177.
21. Grameen Foundation, 2015. Poverty Probability Index. Available at: https://www.povertyindex.org/ [Accessed October 9, 2019].
22. Fraval S, Hammond J, Wichern J, Oosting SJ, de Boer IJM, Teufel N, Lannerstad M, Waha K, Pagella T, Rosenstock TS, Giller KE, Herrero M, Harris D, van Wijk MT. 2019. Making the most of imperfect data: a critical evaluation of standard information collected in cross-sectional farm household surveys. Experimental Agriculture 55 (2), 230-250.
23. Organisation for Economic Co-operation and Development (OECD), & Food and Agricuture Organisation of the UN (FAO) (2017). OECDFAO Agricultural Outlook 20172026.

Data Citations

Upload to Dataverse information!