

Documentation on RHoMIS dataset

Romain Frelat

03-12-2024

Objectives

This document records the steps and choices made to transform the RHoMIS dataset into the format of the farmhousehold Data Platform. It goes together and complements the script file `Get_RHOMIS.R`. The original dataset is an updated version of:

van Wijk, M., Hammond, J., Gorman, L., Adams, S., Ayantunde, A., Baines, D., ... & Yameogo, V. (2020). The Rural Household Multiple Indicator Survey, data from 13,310 farm households in 21 countries. *Scientific Data*, 7(1), 46.

Crop table

The crop table contains 8 columns with information on crop cultivation and usage.

Name	Definition	Unit
hhid	household id	
name	name of the crop	
land_area_ha	land cultivated	hectare
harvest_kg	amount harvested	kg
consumed_kg	amount consumed	kg
sold_kg	amount sold	kg
income_lcu	income from sells	lcu
control_consumed_female	female control over crop consumed	percentage
control_income_female	female control over crop income	percentage

Land area

The land cultivated area was transformed into hectare with the rhomis package.

We estimated the land area per crop from the reported proportion of land cultivated per crop (6 categories: all=90%, most=70%, half=50%, underhalf=20%, little=10%, none=0%). We corrected these percentages when necessary to make sure that the sum of the crop areas was equal to the area of land cultivated (sum of the proportions=100%).

For households without proportions per crop, we assumed that the proportion of crop area is equal to the proportion of quantities harvested.

Crop diversity

We recorded all crops cultivated by farmers. When crops were only listed as grown, without any information on production or usage (not considered as a main crop), we only recorded the hhid and the name in the table, all other columns were filled with NA.

Livestock table

The livestock table contains 3 columns with information on livestock herd

Name	Definition	Unit
hhid	household id	
name	name of the livestock	
n	number of livestock kept	

The livestock are in 16 categories: bees, buffalo, camel, cattle, chicken, dogs, donkeys_horses, duck, fish, goats, guinea_pigs, otherpoultry, oxen, pigs, rabbits, sheep.

Livestock production table

The livestock production table contains 7 columns with information on livestock productions:

Name	Definition	Unit
hhid	household id	
name	name of the livestock	
prod	livestock production	
harvest_kg	amount harvested	kg
consumed_kg	amount consumed	kg
sold_kg	amount sold	kg
income_lcu	income from sells	lcu
control_consumed_female	female control over consumption	percentage
control_income_female	female control over sells	percentage

We consider the production of meat, milk, eggs and honey as well as whole animal sales.

Whole animal sales

For the sales of whole animals, the quantities were given in number of animals. We converted the numbers into kg using the TLU conversion factors (1TLU=250kg). As the animal were sold alive, no quantity was consumed, the quantity harvested equals the quantity sold.

##	alpaca	buffalo	camel	cattle	chicken
##	175.0	175.0	175.0	175.0	2.5
##	donkeys_horses	doves	ducks	geese	goats
##	175.0	2.5	2.5	5.0	25.0
##	guinea_pigs	llama	otherpoultry	oxen	pigs
##	2.5	175.0	12.5	175.0	75.0
##	rabbits	sheep	small_mammal	turkey	
##	5.0	25.0	2.5	5.0	

Household information table

The initial household information table contains information on household composition, off farm activities, and food security.

Name	Definition	Unit
hhid	household id	
country	country of the survey	
year	year of the survey	
gps_lat	latitude in decimal degrees	°N
gps_lon	longitude in decimal degrees	°E
hh_size_members	size of the household in number of persons	
hh_size_mae	size of the household in male adult equivalent	MAE
head_age	age of the household head	
head_gender	gender of the household head	'f' or 'm'
off_farm_lcu	off farm income per year	lcu
off_farm_div	diversity of off farm activities	
hdds	household diet diversity score based on 10 groups	
fies	Food Insecurity Experience Scale based on 8 questions	
foodshortage_count	number of months with food shortage	
foodshortage_months	name of the months with food shortage	
currency_conversion	conversion from local currency to power parity purchase usd	lcu/usd

General information

hhid is made of the variable **id_unique**. This variable is a unique identifier for the household that can connect all tables (crop, livestock, and livestock production).

The GPS coordinates are provided with a rounding at 0.1 decimal degree. The currency power parity purchase conversion factor is provided by the World Bank depending on the country and the year of the survey.

Household size

The household composition is converted by the **rhomis** package. The male adult equivalent are calculated with 5 age-groups: 0-4, 5-10, 11-24, 25-50, 51+ and with the corresponding coefficient per gender and age class.

Off farm income

The household composition is estimated from the **rhomis** package based on the farm production and the estimated percentage of income from off-farm activities.

Food security

We used the food security indicators processed by the **rhomis** package.

Crop and Livestock summary

The household information table also contains 28 columns with summary information from crop and livestock tables.

Name	Definition	Unit
hhid	household id	
land_cultivated_ha	total land cultivated	ha
crop_div	number of crop cultivated	
crop_name	names of crop cultivated	
crop_harvest_kg	total crop harvest	kg
crop_yield_kg_per_ha	crop yield	kg/ha
crop_sold_kg	quantity of crop sold	kg
crop_sold_perc	percentage of quantities of crop sold	%
crop_income_div	number of different crop sold	
crop_income_lcu	total income from crop production	lcu
crop_value_lcu	value of crop produced but not sold	lcu
crop_consumed_kcal	energy value from crop consumed	kcal
livestock_tlu	herd size	tlu
lstk_div	number of livestock species herded	
lstk_name	names of livestock species herded	
lstk_harvest_kg	total livestock product harvested	kg
lstk_sold_kg	quantity of livestock product sold	kg
lstk_sold_perc	percentage of livestock production sold	%
lstk_income_div	number of different livestock products sold	
lstk_income_lcu	total income from livestock production	lcu
lstk_value_lcu	value of livestock production not sold	lcu
lstk_consumed_kcal	energy value from livestock consumed	kcal
farm_div	number of crop and livestock species	
farm_harvest_kg	total farm production	kg
farm_sold_perc_kg	percentage of farm production sold	%
farm_income_div	number of different farm products sold	
farm_income_lcu	total income from farm production	lcu
tot_income_lcu	total income (farm + off farm)	lcu
farm_consumed_kcal	energy value from farm production consumed	kcal
off_farm_perc	percentgae of income from off farm activities	%

All the values are calculated automatically with the function `calc_farm_prod()`. The calculations are simple and summarize, per household, the quantities reported in the crop and livestock tables.

For energy conversion, we used estimates mostly from the FoodData Central of the U.S. Department of Agriculture (<https://fdc.nal.usda.gov/>). Below are the energy conversion factors for the twenty most important crops:

##	banana	beans	bush_beans	cabbage	cassava	groundnut
##	890	1480	1480	250	1600	5500
##	maize	millet	plantain	rice	sesame	sorghum
##	3650	3760	890	3600	5730	3390
##	sugarcane	sweet_potato	wheat	yam		
##	400	770	3400	1180		

GIS information

Based on the GPS coordinates of households, we extracted:

- the Dixon farming system classification for Sub-Saharan Africa Dixon et al. 2021
- the population density estimated by the Gridded Population of the World (GPWv4) and download from geodata package
- the travel time to cities estimated by Nelson et al. 2019 and download from geodata package
- the Koeppen’s Climate Classification from:
Beck, H.E., et al. (2018) “Present and future Köppen-Geiger climate classification maps at 1-km resolution”, *Nature Scientific Data*, 5, 180214 DOI 10.1038/sdata.2018.214

Summary

We kept only *rural* households with at least one crop or one livestock. In total, the dataset is made of 50808 households, with information on 317286 crop cultivated and 96959 livestock species herd. The dataset made of the four tables, together with tlu and energy conversion factors are binded together into a **farmhousehold** object and saved into the file `HHDB_RHOMIS_02122024.rds`.