GHGfromFADN: GHG emissions from FADN data

Table of contents

General infos

GHGfromFADN is an R library to estimate GHG emissions from the main productions of the FADN data (Farm Accountancy Data Network).

Emissions are expressed in kgCO2eq for each production at the farm level. For now, only main field crop productions and no animal products are included.

Technology

This R package was built using R 4.1.1 and depends on R >= 3.5.0

Installation

Step 1: Install the devtools package

```
install.packages("devtools")
```

Step 2: Install GHGfromFADN

```
library(devtools)
```

```
install_github("sarahkleclerc/GHGfromFADN")
```

Step 3: Load GHGfromFADN

```
library(GHGfromFADN)
```

Data available

Three data sets can be loaded from the FADN package:

- `dicovars` FADN variable dictionary based on our data request.
- `ref_country` Table of European country codes (full name, country code, FADN code, Eurostat code)
- `ref_nuts` NUTS region table

Templates

This package contains Markdown templates i.e. pre-filled R Markdown documents for certain common analyses or processes.

To find them (having run `library(GHGfromFADN)` in the console beforehand) in R Studio:

File > New file... > R Markdown... > From Template

- `GHGfromFADN-read-fadn` Fast concatenation of all raw FADN files into one Rdata file.
- `GHGfromFADN-crop-ghg` Prepare data and estimate GHG emissions for a given crop production using GHGfromFADN functions. This provides a good usage example of this library functions.

  A few additional patterns are available. These are scripts from our comparison of the effect of converting to organic farming on GHG emissions:

- `GHGfromFADN-treatment-organic` Create a reference table that states the treatment status for each farm id (e.g. treated = organic farm / control = conventional farm).
- `GHGfromFADN-organic-seqplot-crop` Count farms and plot conversion sequences for a given crop.
- `GHGfromFADN-crop-organic-ghg` Prepare data for the IPW and DID analysis
- `GHGfromFADN-crop-organic-ipw` Cross sectional analysis of the impact of the conversion to organic farming on GHG emissions for given crops among wheat/barley/maize. We use IPW and matching methods.
- `GHGfromFADN-crop-organic-did` Difference-in-differences analysis of the impact of the conversion to organic farming on GHG emissions for given crops among wheat/barley/maize. We use staggered DID algorithm from Callaway and Sant'Anna.

GHG emissions equations

Greenhouse gas emissions are estimated based on IPCC tiers 1 guidelines and emission factors (2006 and 2019 refinements) for most of the emission sources as described in Table 1. The system boundaries are cradle to farm gate: feed production and farming operations are included. Table 1 describes the emission sources included in our estimator for crops. We should add emission sources for animals as well. Most of our data is only available at the farm level and we are interested in the impact at the individual product level. Following LCA principles, we estimate emissions at product level whenever possible (eg. emissions from crop residues are attributed to the relevant crop, emissions from heating are attributed to the relevant animals). For sources of emissions for which an estimate is only possible at a higher level (eg. fertilizers, tractor fuels), the impacts are allocated to the relevant products in proportion to their economic value (International Reference Life Cycle Data System (ILCD) Handbook: General guide for Life Cycle Assessmen, 2010).

Emission sources	FADN data	IPCC equations	Emission factor sources
Direct			
Use of synthetic fertilisers	N fertilisers	Equation 11.1 *	IPCC 2019, table 2A.2
Crop residues management	Crop area, production	Equation 11.1 *	Tables 11.1, 11.2, 11.3 *
Indirect			
Atmospheric deposition	N fertilisers	Equation 11.9 *	Table 11.3 *
Leaching and run-off	N fertilisers	Equation 11.10 *	Table 11.3 *
Synthetic fertilisers production	N fertilisers		Carbon Footprint Calculator (2018)
Energy use (inc. production)	Fuel	Equation 3.3.1 **	Table 3.3.1 **

*IPCC 2006, vol. 4

** IPCC 2006, vol. 2

Our approach is consistent with previous work that adapted the IPCC framework for to FADN data (Coderoni and Eposti, 2015 ; Dabkiene et al, 2020). That said, to the best of our knowledge we are the first to apply it to the European FADN, which is less detailed than the national FADNs, and to several years.

Fertiliser use before 2014 and fuel use had to be approximated from their value. For more details, see the dedicated section.

Inputs from spendings

Fertiliser use

N content of synthetic fertilisers with FADN data are not available in FADN data before 2014.

The FADN survey was not initially designed to conduct environmental impact assessments. Some of the physical data necessary to estimate greenhouse gas emissions are not or were not collected. For a few years now (2014 at the earliest and 2017 at the latest, depending on the country), quantities of mineral N inputs in addition to the expenditure on mineral fertiliser is provided. Before this value was collected for all countries we find many outliers and question whether the units were correctly standardised. In the past, others have tried to derive quantities of inorganic fertilisers from the expenditures. Westbury et al. (2011) approach was to divide expenditures on fertiliser inputs by standard price of fertilisers from the Nix Farm Management Pocketbook. Samson et al. (2012) estimate N fertiliser inputs from the area of each crop and the number of animals in each category based on Agreste and Unifa data. At the time, they did not have access to the volume of nitrogen applied to crops through FADN.

Based on 2018 FADN data of conventional (i.e. not organic nor converting to organic farming in 2018) field crop specialist farms, we fit the structural relationship between N inputs and mineral fertiliser expenditure (Equation 1), using the World Bank Fertiliser Price Index to account for the price variation over time (Equation 2).

$$INUSE\_Q = \alpha SE295 + \epsilon \ (1)$$

In the end we apply Equation 1 to obtain the quantity of fertiliser used in all farms before 2017:

$$ INUSE\_Q = 0.00086815 \left(\frac{\rho_y}{100}\right)^{-1} SE295 + \epsilon \ (2)$$

$y \in [[2004;2019]]$ indicates the year of prediction. $INUSE\_Q$ is the N input in tonnes and SE295 the amount spent on fertilisers inputs expressed in EUR. Finally, $\rho_y \in \mathbb{R}$ is the value of the real price index base 100 in 2018, for the year y.

Energy use

Fuels. Similarly to fertilizers, a conversions from euros to liters is needed for fuel consumption. Fuel prices are obtained from the Weekly Oil Bulletin (Eurostat) which gives details by fuel type, country and with historical depth. We use the gasoline price excluding taxes as it is very similar to off-road diesel prices actually paid by farmers.

Electricity. Electricity is excluded from the emission framework for field crops as it is marginal.

Organic manure emissions

In our framework, emissions from manure are fully allocated to animal products as they are assumed to be waste from production. Manure is commonly given away by livestock farms which supports our "waste" assumption. However, for a few concentrated organic fertilisers such as poultry manure, there is enough demand to create a market, which would then require to attribute manure-related emissions to the cereals on which they are applied. Such an attribution cannot be done with FADN data, as the quantity of manure used is not available.

Usage

See R Markdown Template named GHGfromFADN-crop-ghg

```
# Load libraries  --------------------------------------------------------
library(tidyverse) # general operations
library(data.table) # data
library(GHGfromFADN) # use our library


# Parameters -------------------------------------------------------------
# crops we study
var_crop <- c("CWHTC", "CBRL", "CMZ")
var_crop_text <- c("wheat", "barley", "maize")
var_crop_ha <- paste0(var_crop, "_TA") # area variables (ha)
var_crop_y <- paste0(var_crop, "_PRQ_kg") # production variables (kg)
var_crop_yha <- paste0(var_crop, "_YHA_kg") # production/ha


#### The crop we study (1:wheat, 2:barley or 3:maize) -------------------
i <- 1
crop <- var_crop[i]
crop_text <- var_crop_text[i]
rm(i)


############################ DATA ############################
# Data preparation (general) ---------------------------------------------

# Load data
load("FADN.RData")


#### Field crop specialists subset -------
dfc <-
  data_fadn %>%
  # predict INUSE_Q using the function in farmsty package
  GHGfromFADN::predict_inuse() %>%
  # Shannon crop diversity index using the function in farmsty package
  GHGfromFADN::shannon_crop() %>%
  # Only keep crops specialist farms for now: TF8 == 1
  filter(TF8 == 1)

# data.table
setkey(dfc, id)
unique(dfc, by = c("id", "COUNTRY", "YEAR"))
rm(list=setdiff(ls(), c("dfc", "crop", "crop_text")))


#### Remove outliers (INUSE_Q/UAA) -------
# We use z-score < 3 as a threshold

# Data prep (specific to the crop we study - wheat/barley/maize) ---------------
#### TA>0 -------
# Only farms that have surface of the given crop.
dfc <- dfc %>%
  filter(eval(parse(text = paste0(crop, "_TA", ">0"))))


#### PRQ>0 -------
# Only farms that have production of the given crop.
dfc <- dfc %>%
  filter(eval(parse(text = paste0(crop, "_PRQ", ">0"))))


#### TO_TOC>0 -------
# We allocate emissions from fertiliser use according to the contribution of the
# crop in the overall crops output of the farm. This is done using variables `x_TO_TOC`.
# The issue is that in total output variables `x_TO`, stocks are included and
# therefore sometimes stock of the crop x is lower at the end of the year than
```

```r
# it was in the beginning of the year.
# This explains why we have negative emissions for some farms.
# Later, when we will have the data, we will compute total output excluding stocks.
# For now, we exclude these observations.

dfc <- dfc %>%
  filter(eval(parse(text = paste0(crop, "_TO_TOC", ">0"))))


############################# GHG #############################
# GHG emissions -----------------------------------------------

dfc <- dfc %>%
  # synthetic N input
  GHGfromFADN::ghg_fertilisers(crop = crop) %>%
  # crop residues
  GHGfromFADN::ghg_cresidues(crop = crop) %>%
  # fuels: motors and heating
  GHGfromFADN::ghg_fuels(crop = crop) #%>%
# electricity
# GHGfromFADN::ghg_elec(crop = crop, emission_factor = "EU")

# Total emission for crop i at the farm level
dfc <- dfc %>%
  mutate(ghg_total = rowSums(select(dfc, ghg_ferti_use, ghg_ferti_prod, ghg_cresidues, ghg_fuels_motors), na.rm = TRUE))
# NB: ghg_fuels_heat and ghg_elec excluded

# outcome variables
dfc <- dfc %>%
  mutate(ghg_total_ha = ghg_total / get(paste0(crop, "_TA"),.),
         ghg_total_prq = ghg_total / get(paste0(crop, "_PRQ"),.))

# Save dfc for crop i
saveRDS(dfc, file = paste0("dfc_", crop_text,".rds"))
```