MIND STEP

MODELLING INDIVIDUAL DECISIONS TO SUPPORT THE EUROPEAN POLICIES RELATED TO AGRICULTURE

MIND STEP Project

Data requirements for indicators on European policies related to agriculture and data management

17th March 2021

M. Müller, A. Gocht, S. Neuenfeldt, G. Roerink, J. Helming

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566

https://mind-step.eu/

MIND STEP Methodological Objectives

- to develop a highly modular and customisable suite of **Individual Decision Making (IDM) models** focussing on **behaviour of individual agents** in the agricultural sector to better analyse impacts of policies
- to **develop linkages** between new IDM models and current models used at the European Commission to improve the consistency and to broaden the scope of the analysis of policies
- to develop an **integrated data framework** to support analysis and monitoring of policies related to agriculture
- Improve the **empirical grounding** and **behavioural foundations** (profit maximisation) in current IDM and more aggregated models
- to safeguard the governance and future exploitation of the **MIND STEP model toolbox**

Aim of data works in MIND STEP

- Given the independent existence and continuous changes of databases, MIND STEP aims to design and setup database specific interfaces instead of building "one new big database"
- Bottom-up conceptual data framework that integrates IDM units at farm level, sectors and farming systems at various geographical scales
- It shall enable to:
  - monitor and calculate relevant CAP and SDG indicators:
    - Economic sustainability of farming
    - Provision of ecosystem service
    - …
  - provide data and concepts for simulation models

MIND STEP Data Objectives

- Identify requirements to build a bottom-up conceptual data framework and develop a guide for building standardised data interfaces for the project
- Develop a common data processing plan to share the effort with parallel consortia (BESTMAP, AGRICORE)
- Select, develop and release interfaces to access economic, bio-physical and data of existing models (like GlOBIOM or MAGNET) using state of the art ICT approaches, like REST API, R package distributions, Web Map feature Service)
- Develop, apply methodologies to merge economic (full population and survey data) and biophysical data sets of high spatial and temporal resolution

Standardised Interfaces for:

- Farm Economic Databases (FADN, National Farm Statistics)
  - Translation into a common data structure
  - Identification of data gaps
- Bio-physical databases/large scale data (AgroDataCube)
- Current large scale models (GLOBIOM, MAGNET)

Hardware solutions

- Data storage and processing capacities/Access to a computer cluster & software
- Version control systems and continuous integration

Current activity: A guide/handbook to build an interface

- General definition and meaning of the interface (wrapper versus interface)
- Collect properties to be discussed for each interface (Accessibility, technical implementation, processors, dissemination)
- Propose technical solutions
- Each interface is described in the handbook
- Builds on experiences from previous projects like SEAMLESS and FADNTOOL (https://cordis.europa.eu/project/id/265616/reporting)

Link to Farm Econ. Databases: fadnUtils (1)

an R package to handle FADN data

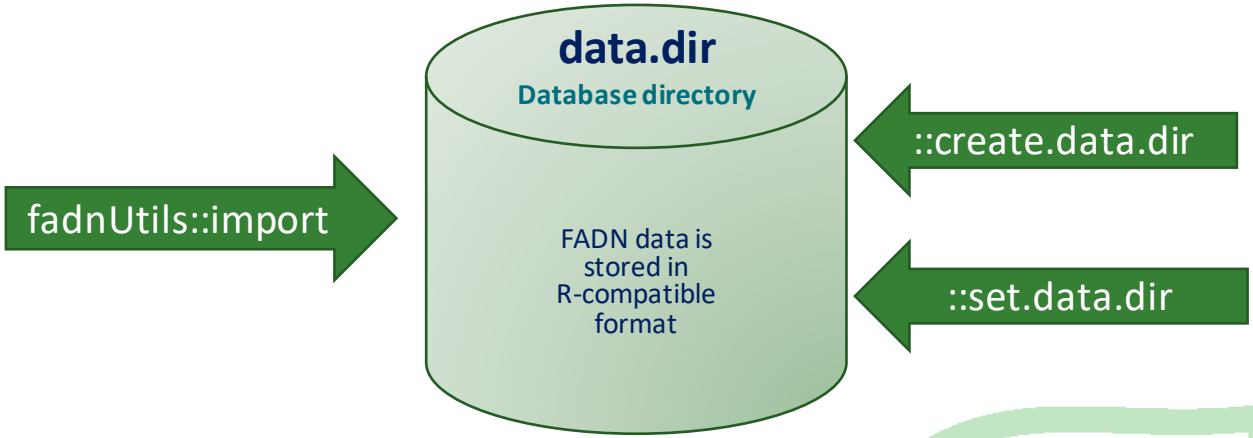
- Developed for preparing the IFM-CAP base-year with 2016 and later FADN data
  (due to significant change in variables from 2014+)
- Sets of functionalities offered
  - Importing and housekeeping FADN data
  - Working with ‘raw’ FADN data
  - Working with ‘calculated’ FADN data

fadnUtils (2)

Importing and Housekeeping FADN data

DG AGRI csv files

Donation ID	Donor ID	Type	Method	Status	Amount	Date
D153	$067	Donation	Credit card	Completed	$100,00	02.05
D154	$123	Shirt	Credit card	Abandoned	$25,00	03.05
D155	$345	Shirt	Paypal	Completed	$25,00	07.05
D156	$367	Donation	Cash	Completed	$50,00	07.05
D157	$121	Shirt	Paypal	Failed	$25,00	07.05
D158	$112	Donation	Credit card	Completed	$75,00	10.05
D159	$055	Donation	Credit card	Completed	$100,00	12.05
D160	$089	Donation	Paypal	Completed	$60,00	13.05
D161	$523	Shirt	Credit card	Failed	$25,00	15.05
D162	$123	Shirt	Cash	Completed	$25,00	17.05
D163	$165	Donation	Paypal	Abandoned	$35,00	17.05
D164	$094	Donation	Paypal	Completed	$75,00	21.05
D165	$054	Donation	Credit card	Completed	$50,00	22.05
D166	$031	Shirt	Credit card	Completed	$25,00	24.05
D167	$451	Donation	Cash	Completed	$100,00	24.05
D168	$112	Shirt	Credit card	Abandoned	$25,00	27.05
D169	$002	Donation	Paypal	Failed	$50,00	27.05
D170	$112	Shirt	Credit card	Completed	$25,00	30.05
Total May					$895,00	

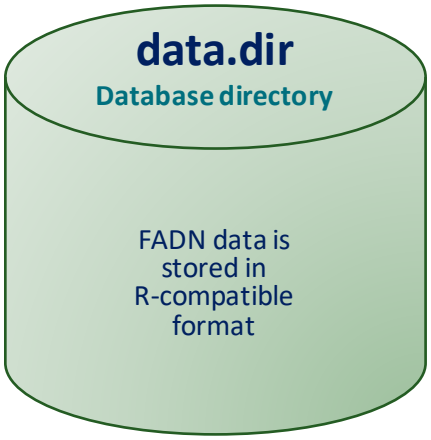



```
> show.data.dir.contents()
{ 'description': 'Phase-0 of IFM-CAP fadn module', 'created-by':'Dimitrios Kremmydas@IFM-CAP-team', 'created-at':'2020-04-21' }

---------------------------------------

Available raw_str_map:
None

---------------------------------------
Available fadn.raw.rds:
    COUNTRY 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016
 1:     BEL    1    1    1    1    1    1    1    1    1    1
 2:     BGR    1    1    1    1    1    1    1    1    1    1
 3:     CYP    1    1    1    1    1    1    1    1    1    1
 4:     CZE    1    1    1    1    1    1    1    1    1    1
 5:     DAN    1    1    1    1    1    1    1    1    1    1
 6:     DEU    1    1    1    1    1    1    1    1    1    1
 7:     ELL    1    1    1    1    1    1    1    1    1    1
 8:     ESP    1    1    1    1    1    1    1    1    1    1
 9:     EST    1    1    1    1    1    1    1    1    1    1
10:     FRA    1    1    1    1    1    1    1    1    1    1
11:     HRV    0    0    0    0    0    0    1    1    1    1
12:     HUN    1    1    1    1    1    1    1    1    1    1
13:     IRE    1    1    1    1    1    1    1    1    1    1
14:     ITA    1    1    1    1    1    1    1    1    1    1
15:     LTU    1    1    1    1    1    1    1    1    1    1
16:     LUX    1    1    1    1    1    1    1    1    1    1
17:     LVA    1    1    1    1    1    1    1    1    1    1
18:     MLT    1    1    1    1    1    1    1    1    1    1
19:     NED    1    1    1    1    1    1    1    1    1    1
20:     OST    1    1    1    1    1    1    1    1    1    1
21:     POL    1    1    1    1    1    1    1    1    1    1
22:     POR    1    1    1    1    1    1    1    1    1    1
23:     ROU    1    1    1    1    1    1    1    1    1    1
24:     SUO    1    1    1    1    1    1    1    1    1    1
25:     SVE    1    1    1    1    1    1    1    1    1    1
26:     SVK    1    1    1    1    1    1    1    1    1    1
27:     SVN    1    1    1    1    1    1    1    1    1    1
    COUNTRY 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016
```

::show.data.dir.contents

fadnUtils (3)

Working with 'raw' FADN data ('raw'=original variables)



fadnUtils::load*

Load combination of countries, years

```
# You can also load for combinations of COUNTRY-YEAR
my.data = load.fadn.raw.rds(
  countries = c("ELL", "ESP"),
  years = c(2014,2015)
)
```

Load subset of variables

```
# You can also load a subset of the columns
#  Here, we only id, nuts2, weight, tf8 and
#  the electricity expenses for all years and countries
my.data = load.fadn.raw.rds(
  col.filter = c("ID","NUTS0","NUTS2", "SYS02","TF8","IELE_V")
)
```

Use metadata on variables

```
> col.codes.subsidies.dt
     code           common.name                                                                    descr            category financed per units
  1: 10000 SAGRPRCTCLIMENVSUBGR_T Agricultural pract. beneficial for climate and environment greening subsidy. Type decoupled.environment                T
  2: 10100         SCRPDVRSUBGR_N                                  Crop diversification greening subsidy. Number decoupled.environment                N
  3: 10100         SCRPDVRSUBGR_T                                    Crop diversification greening subsidy. Type decoupled.environment                T
  4: 10200        SPERMGRSSUBGR_N                                  Permanent grassland greening subsidy. Number decoupled.environment                N
  5: 10200        SPERMGRSSUBGR_T                                    Permanent grassland greening subsidy. Type decoupled.environment                T
 ---
396:  2900           SNO1K2KSUB_V                 Subs not allocated to activity or registered under any other Value          exceptional                V
397:  9000     SDIFPRVSACCYRS_1_V                        Differences from previous accounting years Value with fi 1          exceptional       1        V
398:  9000     SDIFPRVSACCYRS_2_V                        Differences from previous accounting years Value with fi 2          exceptional       2        V
399:  9000     SDIFPRVSACCYRS_3_V                        Differences from previous accounting years Value with fi 3          exceptional       3        V
400:  9000       SDIFPRVSACCYRS_V                                 Differences from previous accounting years Value          exceptional                V
>
```

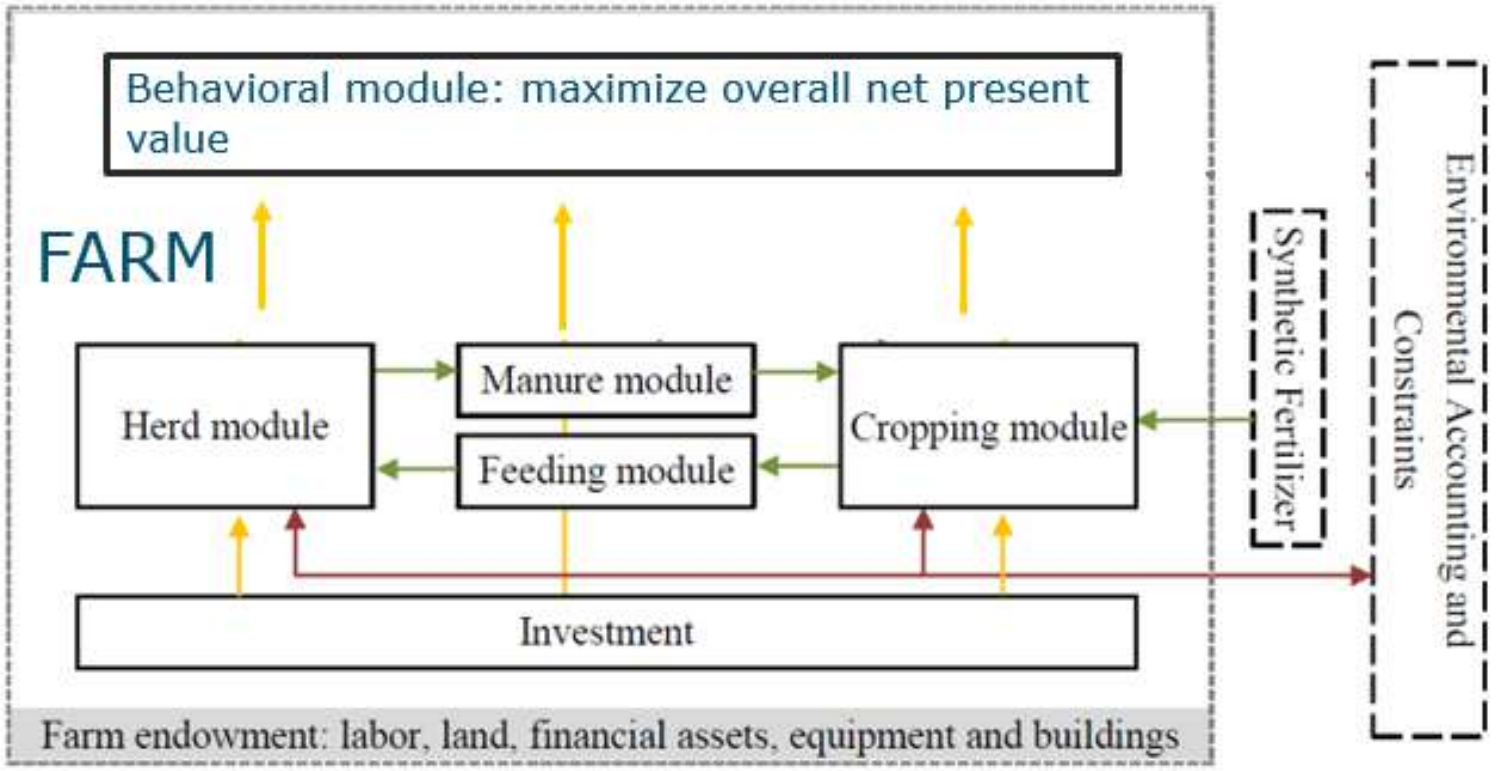
Load subset of farms

```
# You can also load a subset of the rows
#  Here, load only farms that belong to TF8-16, for
#  all years and countries
my.data = load.fadn.raw.rds(
  row.filter = "TF8==16"
)
```

FARMDYN: General description

- Farm level model, originally developed at the University of Bonn
- Flexible, modular template-based, bio-economic model to simulate farms with different branches, currently: dairy, arable, beef cattle, fattening pigs and sows;
- Maximising net-income over large number of years under constraints:
  - Financial-economic (costs, revenues, income, depreciation, investments, taxes, etc.)
  - Agronomical (e.g. feeding, crop rotation, plant nutrition and fertilisation)
  - Policy (EU CAP, environmental policies e.g. manure application)
  - Farm endowments (land, labour, financial assets (liquidity, credits), machinery and buildings)

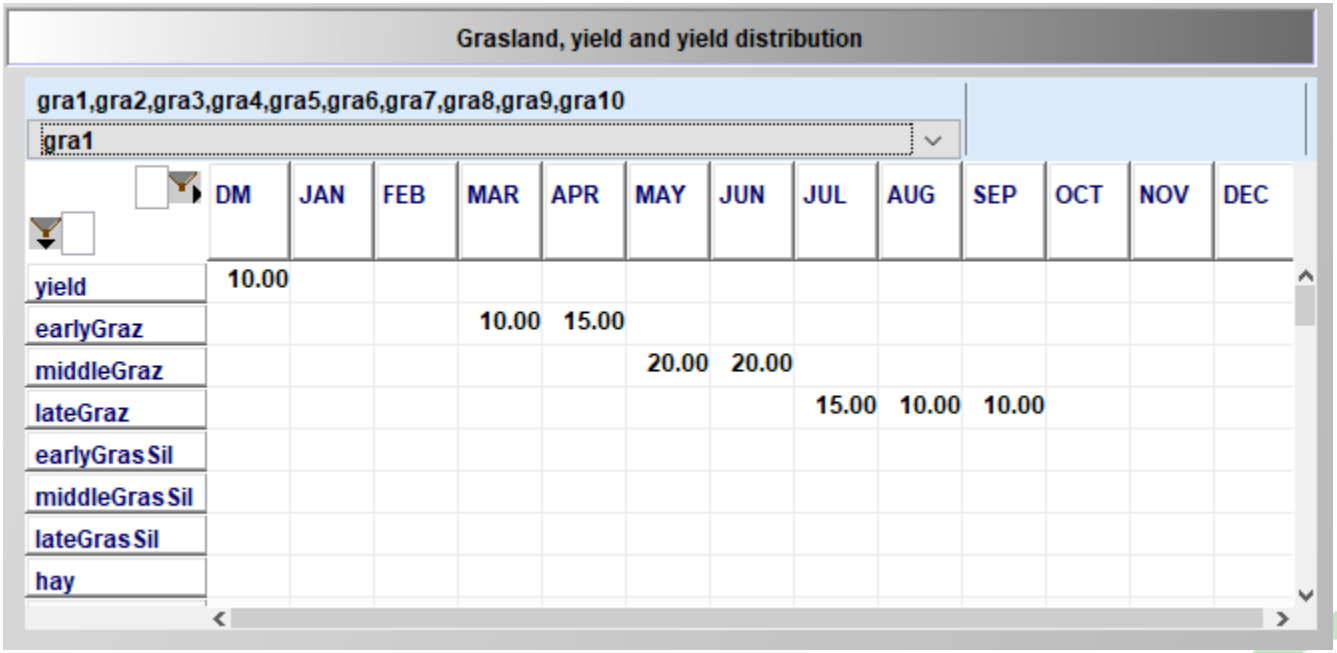
FARMDYN: Schematic



Remark: represents mass transfers from one module to another
represents monetary transfers
represents environmental and related transfers.

Source: Britz et al., 2016

Structure for grassland data in FarmDyn



Grasland, yield and yield distribution

gra1,gra2,gra3,gra4,gra5,gra6,gra7,gra8,gra9,gra10

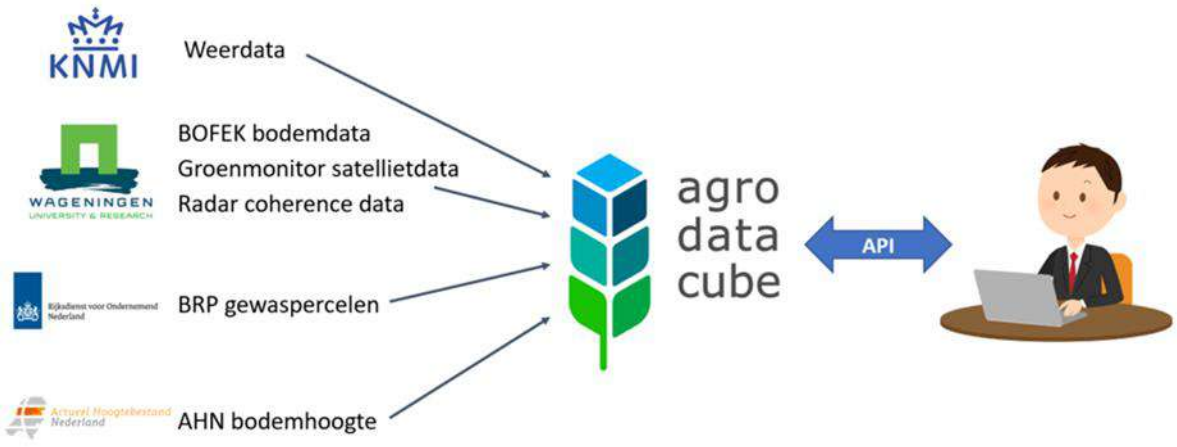
gra1

	DM	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
yield	10.00												
earlyGraz				10.00	15.00								
middleGraz						20.00	20.00						
lateGraz								15.00	10.00	10.00			
earlyGras Sil													
middleGras Sil													
lateGras Sil													
hay													

- Grasslands yields by production system and seasonality generally not available from FADN and national farm statistics
- Solution: populate FarmDyn grassland tables using satellite images

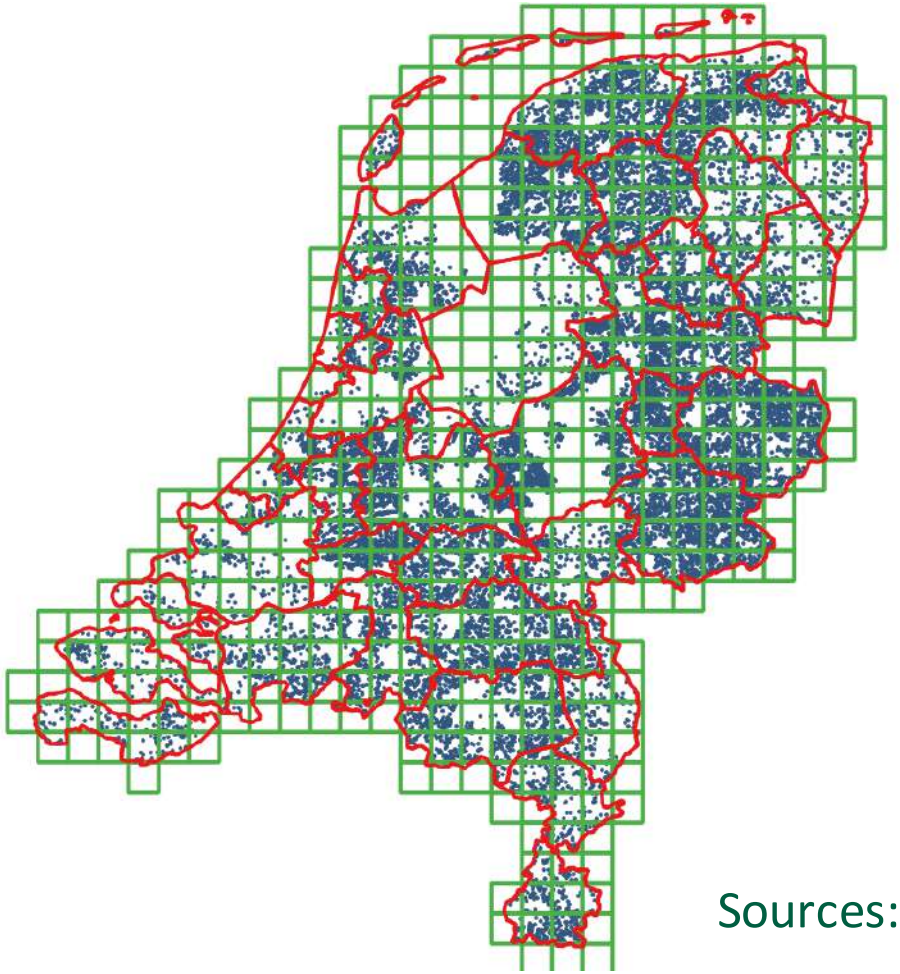
AgrodataCube: Summary

- Big Open Data collection for Agri-Food Applications (Janssen et al 2018)
- Provides a large collection of both open data and derived data for use in agri-food applications
- Open data has been collected from, among others, the Dutch government, Rijkswaterstaat, KNMI and Wageningen University and Research.
- Aims at building on common agri-semantic standards and stimulates the use of open-source data and to exchange open knowledge across the agri-food chain.



LBT: Census of Dutch farms

BIN (National FADN) is subset of LBT

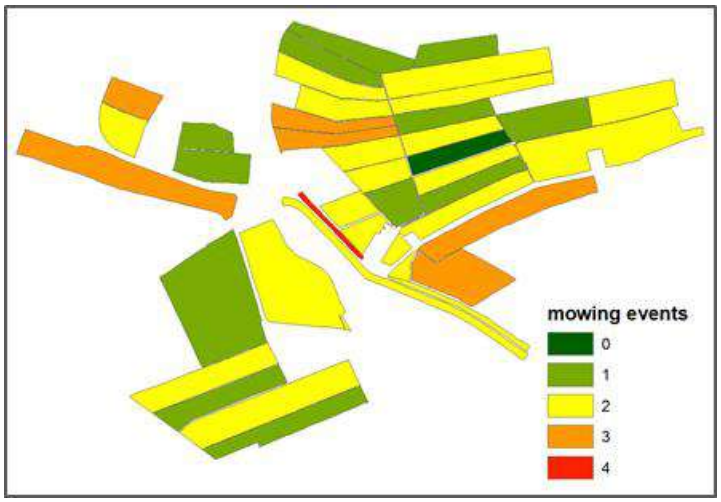


- Groenmonitor.nl

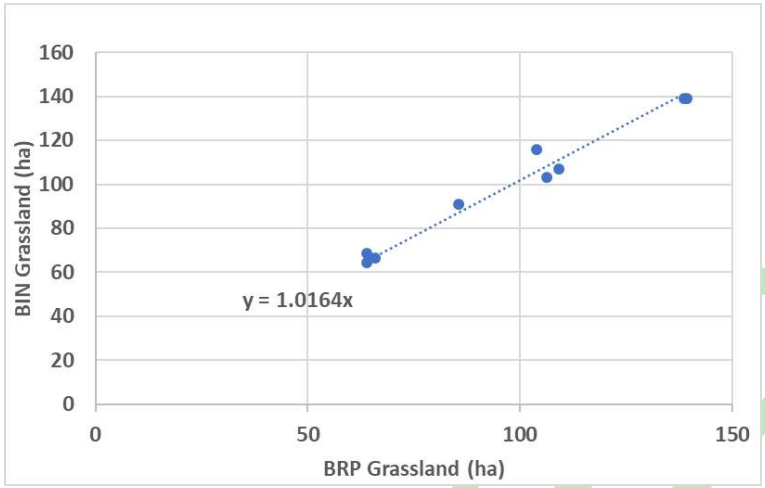
Sources: LBT and AgroDataCube (groenmonitor)

Consistency checks and data extraction

Example: Mowing events



Example: Grassland area from BIN and AgroDataCube (BIN slightly higher)



Ongoing activities

- Identify best way to populate FarmDyn tables (mean DM yield aligned with BIN, nutrient content aligned with literature)
- Compare FarmDyn grassland management results against observations and expert opinion
- Contribute to handbook on data linkages
- Identify more use cases for FarmDyn-AgroDataCube interactions

Transferability to other EU Member States

- Data protection: Parcel-level data, linked to individual farms, are available for NL, but not in all EU member states.
- Satellite images are generally available, but have to be associated with individual farm or spatial groupings
- Publication/exchange of data and results highly restricted to avoid re-identification of farm units.
  - E.g. minimum number of individuals in published units
- This applies also to impact assessment and related indicators!

Probabilistic linking economic statistics (farms) & biophysical data

- Problem
  - Farm-level statistics available at regional level
  - Covering large range of bio-physical conditions
  - High uncertainty for many processes

- Solution
  - Estimating *probability* for each farm in database to be located at a certain *geographic location* with known *environmental characteristics*
    - *Geographic location*: map of spatial units as clusters of 1 km x 1 km grid cells (INSPIRE grid)
    - *Environmental characteristics*: meteorological conditions, soil data, slope, altitude as representative for each spatial unit

Many thanks! Any questions?

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 817566

MIND STEP

https://mind-step.eu/