Exploring precision farming data: a valuable new data source? A first orientation

Tim Punt, Ger Snijkers, Sofie De broe

UNECE Workshop on Statistical Data Collection ‘New sources and New technologies’
14-16 October 2019, Geneva, Switzerland
TechnoBoer heef de toekomst
Three research aim: Fitness for use

Explore the data with respect to:

1. overlap with survey questions
   - Which information does CBS require?
   - What data are available in precision farming
   - Challenges

2. Additional data that is not currently asked
   - Can we use these data for new statistics?

3. Closing the data cycle:
   - Identify Key Performing Indicators (KPIs)
   - Dashboarding

Case study: one arable farmer
Theory!
Does this idea work in practice?
Case study: data from one arable farmer
Precision agriculture cycle

**Winter**
- Draw parcels
- Yield potential
- Tractor lanes

**Spring**
- Fertilization
- Variable planting

**Summer**
- Additional fertilization, pesticides & water
- Based on sensordata

**Autumn**
- Harvest
- Storage
Farmer’s data systems

Input
- Manual
- Machines
- Sensors
- Weather
- GIS data

Monitoring
- Crop Registration Systems

Output
- Ministry of Agriculture
- NSI
- Wholesale Customers
- Surveys

Company Specific Systems

H2S
S2S
Company specific system

DIPPA: Data Integration System for Precision Farming (TUe)
- Farmer’s fields and his actions are central in system
- Log contains 26,000 individual observations
- Data on:
  - Parcels: size, crop, soil conditions, treatments, sowing
  - Crops
  - Crop conditions / crop protection / use of pesticides
  - Activities: when, workload
Data asked for in surveys

1. Combined survey (May)
2. Harvest estimate (October)
3. Crop protection (December)
4. Grassland usage (December)

We tried to complete these questionnaires using the farmer’s data

<table>
<thead>
<tr>
<th>Business area</th>
<th>Harvest (estimate) in kg and ha</th>
<th>Sowing (of winter crops)</th>
<th>Manure</th>
<th>Pesticides</th>
<th>Labour</th>
<th>Subsidies and Grassland</th>
</tr>
</thead>
</table>

CBS Surveys
Overlap precision farming data / questionnaires

- Parcels Area and coordinates
- Harvest in Kg
- Sowing Crops
- Fertilization Products
- Crop Protection Products
- Employees
- Not Available
- (Sensor) data

Company Specific System

- Only one type of crop
- Missing: Fertilisation methods
- Missing: pesticides
- Missing: Gender, age, relation to owner

CBS Surveys

- Business Area
- Harvest Estimate
- Sowing
- Manure
- Pesticides
- Labour
- Subsidies & Grassland Usage
- Not Asked
Overlap precision farming data / questionnaires

Company Specific System
- Parcels Area and coordinates
- Harvest in Kg
- Sowing Cost
- Fertilization Products
- Crop Protection Products
- Employees
- Not Available
- (Sensor) data

Promising!
But many data challenges

Bussiness Area
Harvest Estimate
Sowing
Manure
Pesticides
Labour
Subsidies & Grassland Usage
Not Asked

CBS Surveys
Data challenges for this database

- Lack of metadata
  - Data definitions not always clear
  - Units of measurement are missing (e.g. Kg, ha, Kg/ha)
  - Limited insight on data generation process

- Different data formats
  - Datum, variable names etc.

- Wrong values
  - Freezing temperatures in Augustus
  - Fields in Kazakhstan

- Outliers

- Missings
Data challenges as input for official statistics

‘Fitness for use’ issues:

• Privacy, security, data ownership, data sharing:
  - Trust!
  - Getting access to the data: Who owns the data? Data collection method?

• Quality issues:
  - Sensors as measurement instrument: valid measurements, variance
  - Relevance of the data: correlation with statistical concepts?
  - Unit issue: are the data about the correct unit (fields in Belgium)?

• Ubiquity & standardisation of systems (interoperability):
  - Market penetration: Is the data widespread available in the sector?
  - Interoperability: interface standardisation between systems

• Data Harmonisation:
  - Join data from multiple farmers/sensors into one coherent database?

• Stability of (meta)data delivery in the future
Conclusion: Are (these) sensor data a valuable new source?

• ‘Fitness for use’ issues:
  Combination of primary data and secondary data characteristics:
  - NSIs are not in control of the data: found data vs. designed data
  - Similarities with admin data
  - Market penetration and accessing the data

• Developments are going quickly:
  - We need to be involved now!
  - More studies are needed

• Short term: crop registration systems

Yes, valuable and promising, but still a long way to go ...

Developments and Experiences in other countries?
Farmer’s data systems

Input

- Manual
- Machines
- Sensors
- Weather
- GIS data

Monitoring

- crop Registration Systems
- Company Specific Systems

Output

- Ministry of Agriculture
- NSI
- Wholesale Customers

Surveys

S2S

H2S

14
Overlap crop registration system / questionnaires
Potential

Harvest estimate: 2

Harvest: 1
Sowing wintercrops: 1
Labour & biological farming: 1
Area: 1
Manure: 3
Subsidies: 5
Mushrooms, bulbs & chicory: 1
Farm animals and manure: 8
Combined assignment: 22

Crop protection: 3
Grassland: 2
Horticulture: 3
Use: 1
Hay: 1

General: 1
Chemical: 1
Mechanical: 1

Manual: 7
"Harvest registration systems"/DIPPA: 10
NA: 12

16
ARE THERE ANY QUESTIONS? FEEL FREE TO ASK ANYTHING AT ALL.
Facts that matter