EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS

2024 ESTORIL - PORTUGAL

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

An innovative approach to using a variety of data sources and statistical methods for micro-level statistics collection

Ylva Olsson, Swedish Board of Agriculture
Ylva.Olsson@jordbruksverket.se
Anders Grönvall, Swedish Board of Agriculture
Jesper Fransson, Swedish Board of Agriculture
Elin Lund, Swedish Board of Agriculture

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

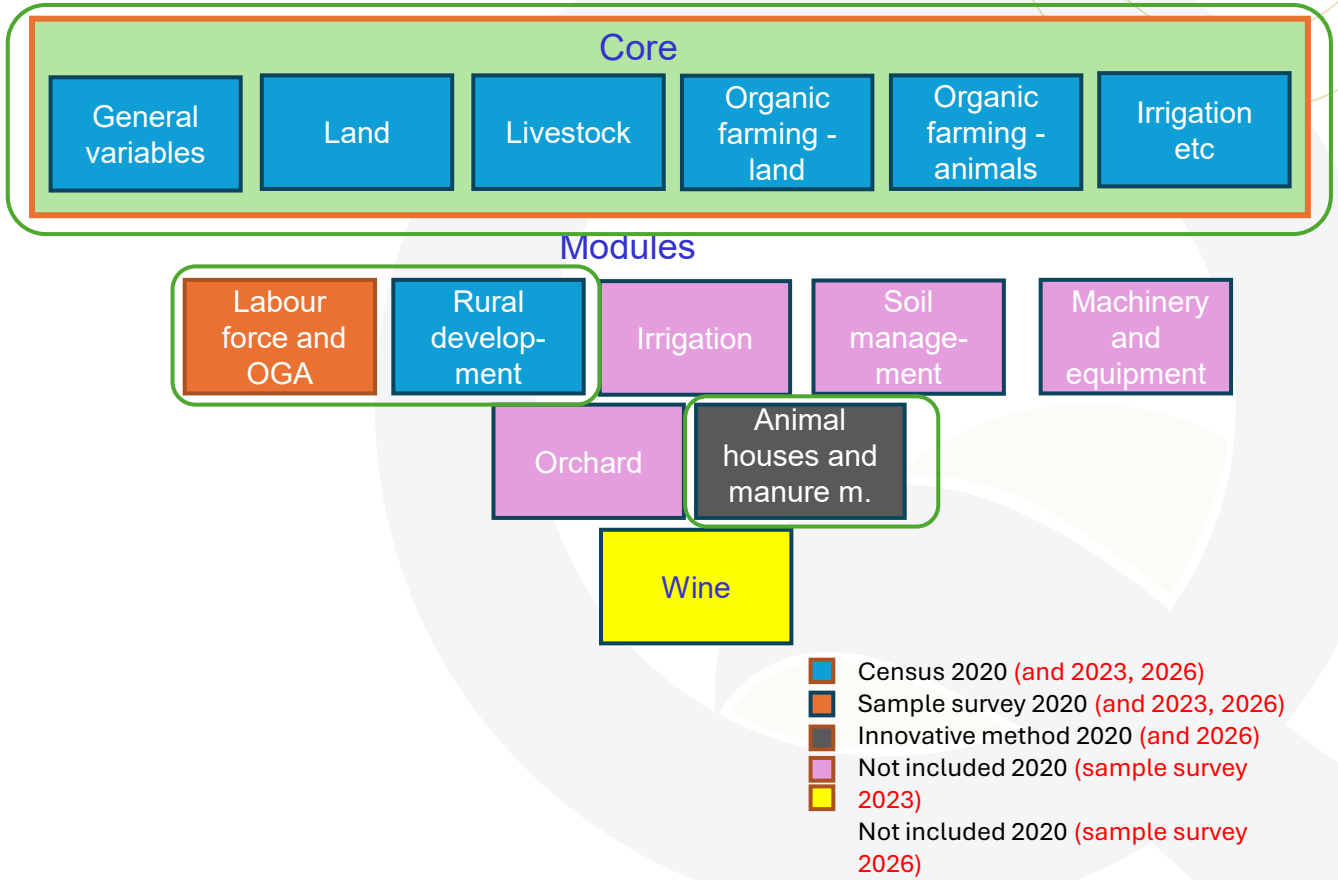
eurostat

The conference is partly financed by the European Union

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL
eurostat
The conference is partly financed by the European Union

Integrated farm statistics (IFS), 2020

Regulation (EU) 2018/1091 of the European Parliament and of the Council of 18 July 2018 on **i**ntegrated **f**arm **s**tatistics and repealing Regulations (EC) No 1166/2008 and (EU) No 1337/2011



The conference is partly financed by the European Union

Integrated farm statistics (IFS) - regulation

From IFS-regulation

Article 4

Data sources and methods

1. For the purpose of obtaining the data referred to in this Regulation, Member States shall use one or more of the following sources or methods, provided that the information allows for the production of statistics that meet the quality requirements laid down in Article 11:

(a) statistical surveys;

(b) the administrative data sources specified in paragraph 2 of this Article;

(c) other sources, methods or innovative approaches.

2. Member States may use information from the integrated administration and control system (IACS) established by Regulation (EU) No 1307/2013 of the European Parliament and of the Council ([2]), the system for the identification and registration of bovine animals established by Regulation (EC) No 1760/2000 of the European Parliament and of the Council ([3]) and the system for the identification and registration of ovine and caprine animals established by Council Regulation (EC) No 21/2004 ([4]), the vineyard register implemented in accordance with Article 145 of Regulation (EU) No 1308/2013 of the European Parliament and of the Council ([5]), and the organic farming registers set up pursuant to Council Regulation (EC) No 834/2007 ([6]). Member States may also use administrative sources associated with specific rural development measures.

3. Member States which decide to use the sources, methods or innovative approaches referred to in point (c) of paragraph 1 shall inform the Commission (Eurostat) during the year preceding the reference year and shall provide details concerning the quality of the data obtained from that source, method or innovative approach and the data collection methods to be used.

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL
eurostat
The conference is partly financed by the European Union

Integrated farm statistics (IFS)

From Handbook

2.3.3 Other sources

Member States which decide to use the sources, methods or innovative approaches referred to in Article 4 (1) (c) of Regulation (EU) 2018/1091 (other sources, methods or innovative approaches) shall inform the Commission (Eurostat) during the year preceding the survey reference year and shall provide details concerning the quality of the data obtained from that source, method or innovative approach and the data collection methods to be used.

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL
eurostat
The conference is partly financed by the European Union

Animal housing and manure management – Background

Information on animal stables and fertilisers has been collected at various times in connection with structure surveys for agriculture, for example in 2010 and 2016:

- High partial non-response
- Difficult questions in survey form – uncertain quality
- Risk of it affecting the response frequency of other parts
- For national needs, Statistics Sweden conducts a fertiliser and manure survey every three years
- A lot of information is available in different places – however, not collected in a register
  - Information from different advisory organisations
  - All new animal houses have to be reported to and approved by the county board
  - Every third year Statistics Sweden conduct a survey on manure
  - In 2010 and 2016 there where parts of this area in the structure survey for agriculture
  - The cattle register and then production place register have lot of information

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL
eurostat
The conference is partly financed by the European Union

Animal housing and manure management – modelling

Modelling – Whats is it?

How can we, based on existing information, create a model that satisfactorily generates an expected result for each variable for the holding.

- Direct or indirect imputation based on:
  - distribution
  - legal or practical conditions in real life
  - relationships and correlations that are more or less known
- To get x, we can ask about y instead
- Use existing registers
- Use information from other surveys
- Regression models between several variables
- Combination of different solutions
- Regional conditions
- Rules that govern, both on national and regional level and what is not allowed in Sweden (we assume that every holding follows the regulation)

The conference is partly financed by the European Union

Modelling

		Regulations							from IFS (census)													sources							Modelling				
		organic		Store manure / region		outdoor / region		…	number of animals						crops (hectare)						Calculations			register			survey			…	variabel / variabels				
holding	region	land	animal	regions	LU	animal	days	…	dairy cow	other cow	calves	…	pigs	horses	wheat	green fodder	…	fallow land	under glass	pasture / meadow	livestook unit (LU)	manure production	…	A	B	…	X	Y	…	…	a	b	c	d	…
1	A	yes	yes	**Yes**	if over 10	if bovine	90		10		30			10	20					10														
2	B		yes	No		if bovine	60						200																					
3	A	no	no	**Yes**	if over 10	if bovine	90			100	50									200														
4	A	yes		**Yes**	if over 10	if bovine	90								40			10																
5	B	no	no	No		if bovine	60			50			100			250	50			20														
6	B	no		No		if bovine	60												200															
7	C	yes		**Yes**	if over 100	if bovine	120									20		10		10														
8	B		no	No		if bovine	60		100		200									50														
9	C	yes	yes	**Yes**	if over 100	if bovine	120						50		50				30	10														
10	A	no	no	**Yes**		if bovine	90			10	10			5		10	10																	
9	C	yes	yes	**Yes**	if over 100	if bovine	120						50		50				30	10														
10	A	no	no	**Yes**		if bovine	90			10	10			5		10		10																

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL
eurostat
The conference is partly financed by the European Union

Create logic

For those data we had we made microdata imputation on each holding.

We had to create some kind of logical solutions for:

- What source to rely on in different situations
  - Inconsistent information in different sources
- Different situations/sources have different logical solutions
- Different sources in certain combination give a full view
- Some situations are incompatible – how to deal with this

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL
eurostat
The conference is partly financed by the European Union

Create logic

"Empty cells" imputed by:

Stratify the population based on different information for different variable. For those with missing values put information from other holding in same strata. Different methods depending on type of missing data.

		Regulations							from IFS (census)														Calculations			sources							Modelling				
		organic		Store manure / region		outdoor / region		…	number of animals						crops (hectare)											register			survey			…	variabel / variabels				
holding	region	land	animal	regions	LU	animal	days	…	dairy cow	other cow	calves	…	pigs	horses	wheat	green fodder	…	fallow land	under glass	pasture / meadow	livestock unit (LU)	manure production	…	A	B	…	X	Y	…	…	a	b	c	d	…
1	A	yes	yes	**Yes**	if over 10	if bovine	90		10		30			10	20					10															
2	B		yes	No		if bovine	60						200																						
3	A	no	no	**Yes**	if over 10	if bovine	90			100	50									200															
4	A	yes		**Yes**	if over 10	if bovine	90								40			10																	
5	B	no	no	No		if bovine	60			50			100			250	50			20															
6	B	no		No		if bovine	60												200																
7	C	yes		**Yes**	if over 100	if bovine	120									20		10		10															
8	B		no	No		if bovine	60		100		200									50															
9	C	yes	yes	**Yes**	if over 100	if bovine	120						50		50				30	10															
10	A	no	no	**Yes**		if bovine	90			10	10			5		10	10																		

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL
eurostat
The conference is partly financed by the European Union

Methods

In practice we have “invented” a method for each variable. We have used methods like:

- Imputation from registers
- Used statistical distribution in different homogeneous groups
- Randomly select data from similar holding
- Strong relationships with other variables
- Used factors from various sources
- Assuming holding follow legal rules
- Adapted based on regional conditions and rules
- "Consequential effects" of a previous imputation
- Assumptions and some generalisations

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL
eurostat
The conference is partly financed by the European Union

Problems

- To understand the subject
- Different sources provide different information. How to evaluate the different sources?
- We would like to have access to more sources
  - It is difficult to explain in advance why/how
  - Legal issues
  - It could take time to obtain the data from other sources depending on how easy they can extract data from their systems
- How to assess quality
- Link data from different sources

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL
eurostat
The conference is partly financed by the European Union

Conclusions / Lessons learned

Conclusions

- Less response burden for the holding
- We have information for each holding in the survey, some information of better quality and some of worse quality. In total we think the results are good enough for each holding and the data passes the validation rules that Eurostat have defined. We think that aggregated data have good quality, but it is not necessary that data for a single holding is correct.

Lessons learned

- A well known problem is that information can be linked to the holding in different ways.
- Extremely knowledge – oriented
- It takes time
- It is always possible to do more
- More preparatory work
  - The models could be finished before the survey
  - Identify if there are one or few individual variables that in some way could be collected in connection with the survey (in this case "Other organic and waste-based fertilisers than stable manure used on the farm").
- Properly record which methods and models were used for which variables in the documentation. Researchers who wish to use data for all holdings in the survey must know that the methods used and the quality of the data are not always exactly the same even for a given variable.

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL
eurostat
The conference is partly financed by the European Union

Quality framework

New data sources and methods require new quality frameworks.

In the system of IFS-statistics we are by regulation allowed alternative methods for creating statistics but the quality reporting must be fitted into a fixed template which is largely based on the results being derived from statistical surveys or to some extent administrative registers.

In this case the total IFS-survey consists of data from survey, data from administrative registers but also data from modelling (innovative methods). Existing quality frameworks do not take this into account and with fixed templates to report the quality, the problem becomes even greater.

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Thank you

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly
financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS

2024 ESTORIL - PORTUGAL

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union