

# Improving efficiency of the sample design in the Finnish horticultural survey

- Impact of introducing Standard Output as a threshold and imputation methods

Anna-Kaisa Jaakkonen,  
Mika Kuoppa-aho and  
Johanna Laiho-Kauranne

**Luke - Natural Resources Institute Finland**

BanoCoss Conference, Helsinki, 28.8.2015

# Horticultural statistics is part of ESS

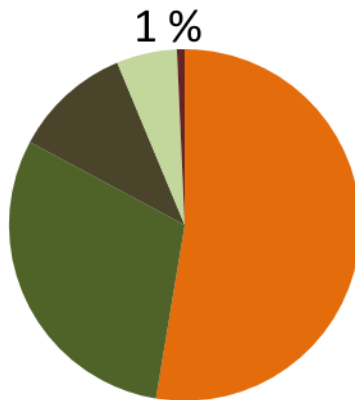
- The horticultural statistics illustrate:
  - the structure of commercial horticultural enterprises,
  - their outdoor and greenhouse production & mushroom cultivation.
- Statistical category: Agriculture, Forestry and Fishery
- **Official Statistics of Finland (OSF) & European Statistical System (ESS):**
  - **Regulation (EC) No 543/2009 of the European Parliament and of the Council** of 18 June 2009 concerning crop statistics and repealing Council Regulations (EEC) No 837/90 and (EEC) No 959/93 (1)
  - **Regulation (EU) No 1337/2011 of the European Parliament and of the Council** of 13 December 2011 concerning European statistics on permanent crops and repealing Council Regulation (EEC) No 357/79 and Directive 2001/109/EC of the European Parliament and of the Council.
- The data is required from each EU-MS to monitor development of European agriculture for the purposes of CAP – Common Agricultural Policy, which makes 40 % of the total budget of the European Union



# Horticultural enterprises have a significant share of agriculture in Finland

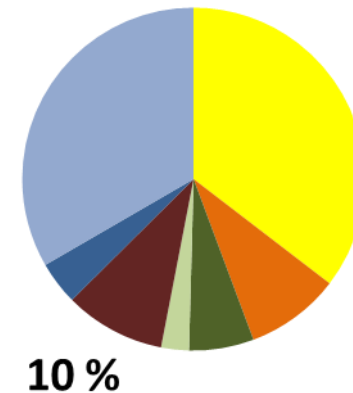
- Horticultural production in 3 500 enterprises, total UUA area 16 300 ha, of which green houses 390 ha
- Share of all agricultural enterprises is 6%, of utilised agricultural area 1 % while the share of market value of economic output is 10 %

Utilized agricultural area in 2014



■ Cereals ■ Grasslands ■ Fallow ■ Others ■ Horticulture

Total calculation of agriculture at current prices



■ Livestock ■ Cereals ■ Grass ■ Other plants  
■ Horticulture ■ Other return ■ Support payments

# Horticultural statistics in Finland

- Total annual sample survey - census of horticultural enterprises since 1984
- Introduction sampling methods has not been solved as both the level of detail required by the EU regulation and the heterogeneity of the enterprises has been very high (details of permanent crops by density & age categories, location, large and small enterprises)
- Large variety of plants & crops: 12 berries & fruits, 33 outdoor vegetables, nurseries, outdoor cut flowers, cultivated mushroom, 17 indoor vegetables, small plant production, cut flowers and 35 different ornamental plants.
- Every third year additional combined survey on the use of energy.
- Mixed mode of advance letter & web survey + telephone interviews
- Web response rate 55 % of all respondents
- Total response rate is close to 98% of horticultural farms
- Coverage of almost 100 % of the land in horticultural production





# The need to modernise and improve the efficiency

- Taking into concern the national information needs & energy statistics
  - including the required volume of production for the European farm structure survey
  - usage of energy
- Preconditions for modernization:
  - To reduce the response burden of the enterprises
  - To reduce the survey costs without sacrificing the accuracy and quality of the survey information
- The importance of the economic value has risen in connection with the monitoring the volumes of production

## Pre-conditions & constraints:

- The key variables to be estimated are totals & distribution by geographic classification
- The units and the areas of production are known from register



# Balance of finding savings

- Using thresholds to reduction of sample size maintains annual survey is a balanced solution to gain savings and reduction of burden
- EU-Regulation on permanent crops:
  - *To avoid placing an unnecessary burden on farms and administrations, thresholds should be established that exclude non-relevant entities from the basic entities in respect of which statistics on permanent crops are to be collected.*
- Other means for savings are:
  - Cut down drastically information contents → reduction of relevance for national data needs – possibly increasing the need for other national data collection
  - From annual to semi-annual using estimation → would increase salary costs of the researchers i.e. 0 savings
  - Abandon mailing of advance letters → reduction in web-response rates
  - Increase web-response by all possible means = reduction in interviewer costs
  - Accept only web-responses → reduction in response rates; and inadequate coverage of the survey:
    - *Statistics to be provided on permanent crops referred to under points (a) to (l) of Article 1(1) shall be representative of at least 95 % of the total planted area producing entirely or mainly for the market of each permanent crop referred to in each Member State.*

# Why the reduction of response burden matter?



Installation of a farmer of 10,7 meters a farmer needs to fill in annually  
– with additional forms for annual leave it makes upto 25 meters of forms  
*excluding agricultural statistics questionnaires and the academic and market research*

Installation has been made by agricultural advisor Mari Tabell in Kiurujoki



# Definition of Standard Output

- **SO i.e. Standard Output** is the average economic yield for farm products per hectar or per production animal excluding agricultural benefits for production.
- **At present threshold is 2 000 eur** in the frame population in statistical register of farms & horticultural enterprises which is based on IACS

	SO eur/ha
Open field vegetables	11 399
Open field berries	4 874
Ornamentals in green houses	631 439
Vegetables in green house	647 148
Oat	499
Potatoes	4 421

*Threshold for EU-subsidies on green houses is 300 m<sup>2</sup>*



# Sensitivity analysis: Impact on estimated production volumes by SO

	SO: 2 000 eur			SO: 6 000 eur			SO: 10 000 eur			SO: 15 000 eur		
	# farms	ha	1 000 kg	# farms	ha	1 000 kg	# farms	ha	1 000 kg	# farms	ha	1 000 kg
Strawberries	1 106	3 298	12 858	1 105	3 298	12 858	839	3 212	12 697	738	3 144	12 569
Highbush blueberry	147	72	115	117	59	102	94	47	82	80	38	76
<b>All berries</b>	<b>1 612</b>	<b>5 797</b>	<b>15 575</b>	<b>1 251</b>	<b>5 608</b>	<b>15 422</b>	<b>1 069</b>	<b>5 445</b>	<b>15 255</b>	<b>928</b>	<b>5 258</b>	<b>15 016</b>
Carrots	361	1 652	74 221	330	1 648	74 164	309	1 645	74 111	284	1 637	73 944
Garlic	67	19	49	52	16	47	44	14	44	38	13	38
<b>All open land vegetables</b>	<b>1 408</b>	<b>9 142</b>	<b>185 284</b>	<b>1 251</b>	<b>9 103</b>	<b>185 284</b>	<b>1 157</b>	<b>9 058</b>	<b>184 977</b>	<b>1 058</b>	<b>8 977</b>	<b>184 280</b>
Tomatoes	332	1 040	39 890	331	1 040	39 889	330	1 040	39 889	330	1 040	39 889
Butter-head lettuce	47	128	2 601	47	128	2 601	47	128	2 601	47	128	2 601
<b>All greenhouse vegetables</b>	<b>566</b>	<b>2 250</b>	<b>83 058</b>	<b>565</b>	<b>2 250</b>	<b>83 058</b>	<b>563</b>	<b>2 249</b>	<b>83 058</b>	<b>563</b>	<b>2 249</b>	<b>83 058</b>
Bedding plants (# 1 000 )	<b>514</b>		<b>38 853</b>	<b>514</b>		<b>38 853</b>	<b>512</b>		<b>38 851</b>	<b>509</b>		<b>38 845</b>

# Sensitivity analysis: Impact of increasing the threshold from 2 000 eur to 10 000 euros, %

	SO: 6 000 eur			SO: 10 000 eur			SO: 15 000 eur		
	# farms	ha	1 000 kg	# farms	ha	1 000 kg	# farms	ha	1 000 kg
Strawberries	-0,1	0,0	0,0	-24,1	-2,6	-1,2	-33,3	-4,7	-2,2
Highbush blueberry	-20,4	-18,3	-11,3	-36,1	-35,3	-28,3	-45,6	-46,7	-33,8
<b>All berries</b>	<b>-22,4</b>	<b>-3,3</b>	<b>-1,0</b>	<b>-33,7</b>	<b>-6,1</b>	<b>-2,1</b>	<b>-42,4</b>	<b>-9,3</b>	<b>-3,6</b>
Carrots	-8,6	-0,2	-0,1	-14,4	-0,4	-0,1	-21,3	-0,9	-0,4
Garlic	-22,4	-15,8	-4,1	-34,3	-26,3	-10,2	-43,3	-31,6	-22,4
<b>All open land vegetables</b>	<b>-11,2</b>	<b>-0,4</b>	<b>0,0</b>	<b>-17,8</b>	<b>-0,9</b>	<b>-0,2</b>	<b>-24,9</b>	<b>-1,8</b>	<b>-0,5</b>
Tomatoes	-0,3	0,0	0,0	-0,6	0,0	0,0	-0,6	0,0	0,0
Butter-head lettuce	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
<b>All greenhouse vegetables</b>	<b>-0,2</b>	<b>0,0</b>	<b>0,0</b>	<b>-0,5</b>	<b>0,0</b>	<b>0,0</b>	<b>-0,5</b>	<b>0,0</b>	<b>0,0</b>
<b>Bedding plants (# 1 000)</b>	<b>0,0</b>		<b>0,0</b>	<b>-0,4</b>		<b>0,0</b>	<b>-1,0</b>		<b>0,0</b>

- The impact is large in production sectors with small economic value and non-existent in greenhouse production, which are larger professional enterprises.
- The focus in horticultural & agricultural statistics is towards to the large production that enters to the market and not in small kitchen farms and less in social aspects

# SO 10 000 eur: impact for production totals is minimal -> imputation for adequate estimates

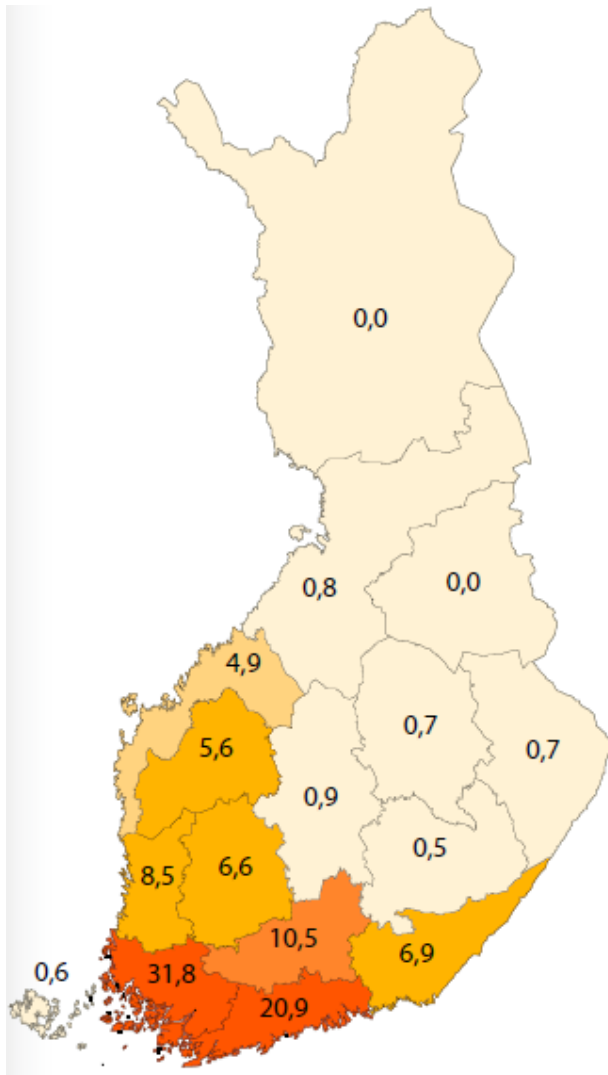
	SO: 2 000 eur			Diff: SO 2 000 eur/ 10 000 eur			% Diff 2 000 / 10 000 eur		
	# farms	ha	1 000 kg	# farms	ha	1 000 kg	# farms	ha	1 000 kg
Strawberries	1 106	3 298	12 858	-267	-86	-161	-24	-3	-1
Highbush blueberry	147	72	115	-53	-25	-33	-36	-35	-28
<b>All berries</b>	<b>1 612</b>	<b>5 797</b>	<b>15 575</b>	<b>-543</b>	<b>-352</b>	<b>-320</b>	<b>-34</b>	<b>-6</b>	<b>-2</b>
				0	0	0			
Carrots	361	1 652	74 221	-52	-7	-110	-14	0	0
Garlic	67	19	49	-23	-5	-5	-34	-26	-10
<b>All open land vegetables</b>	<b>1 408</b>	<b>9 142</b>	<b>185 284</b>	<b>-251</b>	<b>-84</b>	<b>-307</b>	<b>-18</b>	<b>-1</b>	<b>0</b>
				0	0	0			
Tomatoes	332	1 040	39 890	-2	0	-1	-1	0	0
Butter-head lettuce	47	128	2 601	0	0	0	0	0	0
<b>All greenhouse vegetables</b>	<b>566</b>	<b>2 250</b>	<b>83 058</b>	<b>-3</b>	<b>0</b>	<b>-1</b>	<b>-1</b>	<b>0</b>	<b>0</b>
<b>Bedding plants (# 1 000 )</b>	<b>514</b>		<b>38 853</b>	<b>-2</b>	<b>0</b>	<b>-2</b>	<b>0</b>		<b>0</b>

Proportional difference is high for plants with small impact on # of farms & economical output & volumes

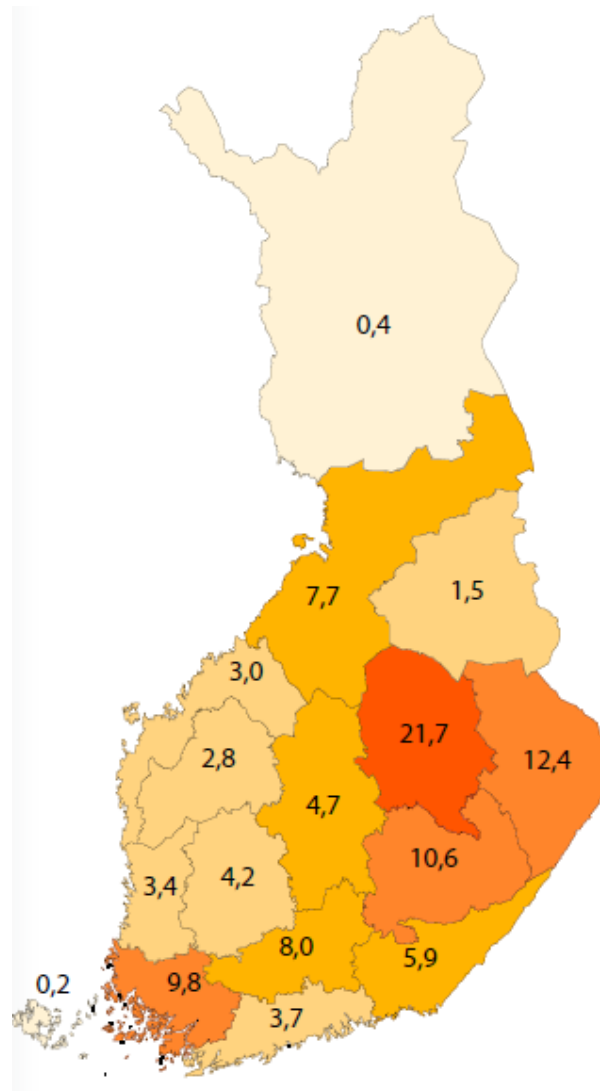
Proportional difference is non-significant for larger production branches

If the threshold is risen – the impact can be corrected with imputation

# Horticultural farms have huge regional differences (2013)



Horticultural farms mainly producing vegetables on open fields, % by ELY



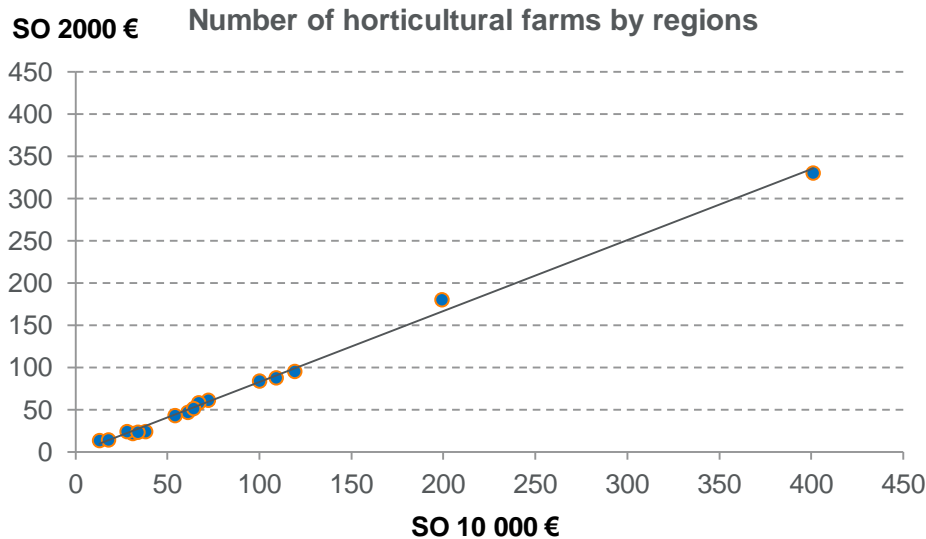
Horticultural farms mainly producing berries, % by ELY

© Luonnonvarakeskus

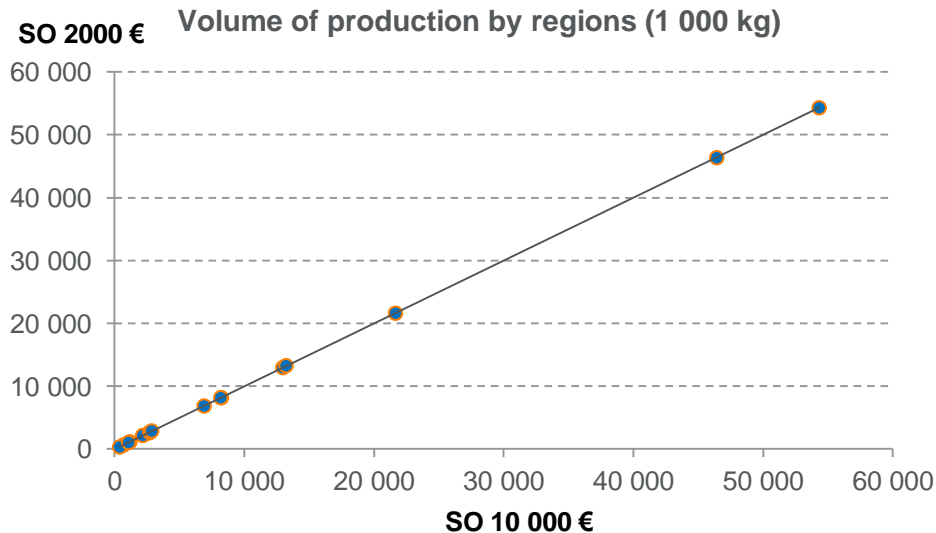


# Studying the impact of increasing the threshold by regions

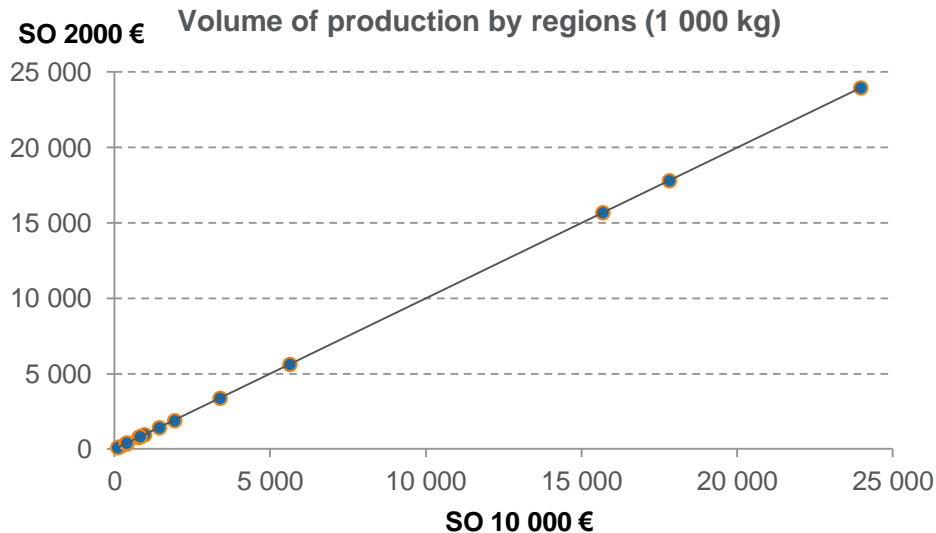
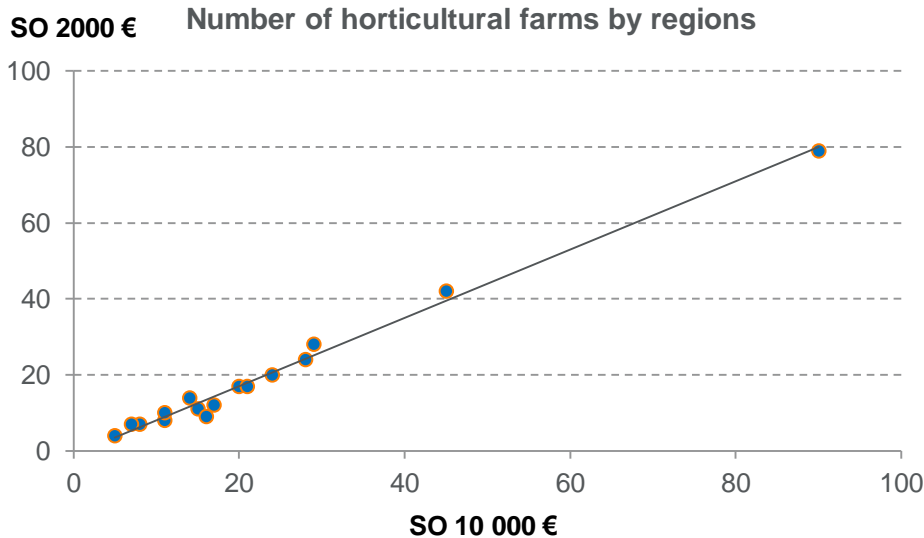
- horticultural farms of main production is in vegetables in open land



Overall impact:  
-17,71 % of farms  
-0,23 % of production



# Studying the impact of increasing the threshold by regions - horticultural farms producing mainly carrots in open land

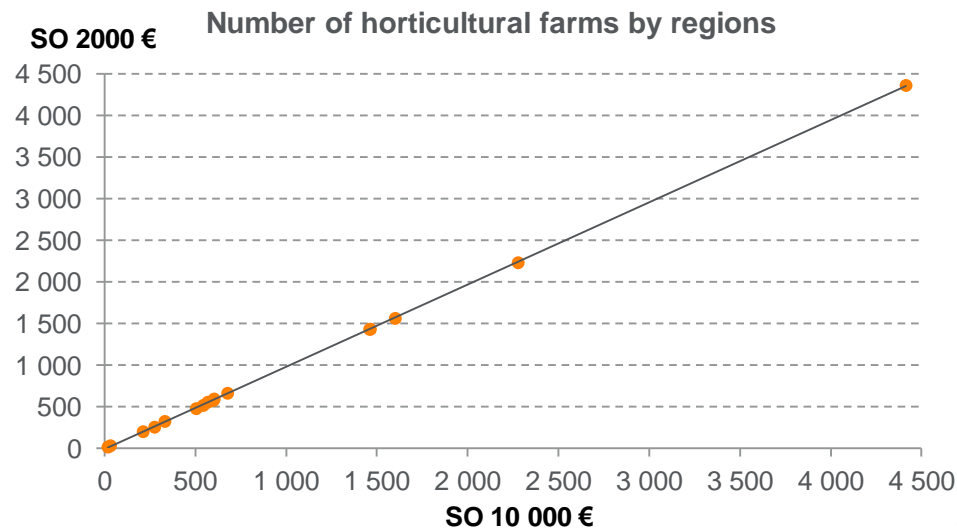
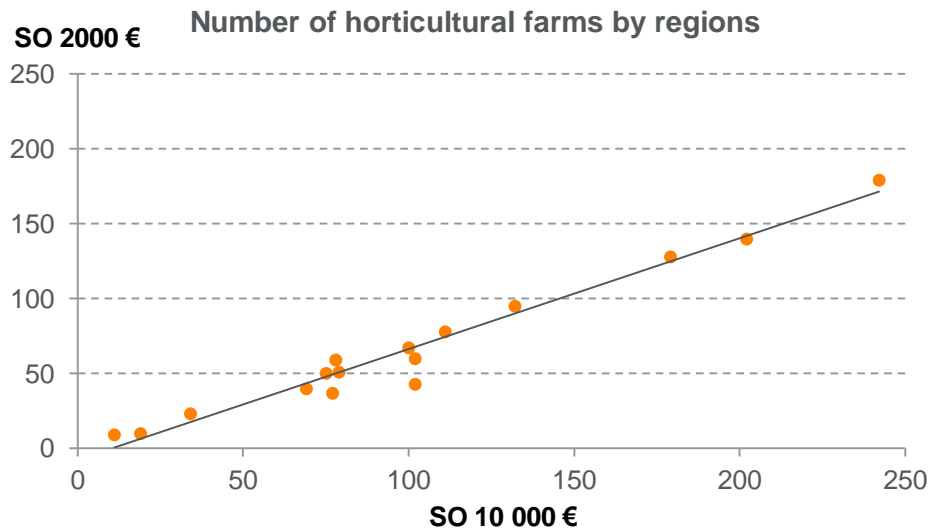


## Overall impact:

- 14,40 % of farms
- 0,15 % of production

But deviation  
by regions increases  
when analysing bias  
in more detailed  
classification

# Studying the impact of increasing the threshold by regions - horticultural farms producing mainly berries



## Overall impact:

- 33,68 % of farms
- 2,05 % of production

But deviation  
by regions increases  
when analysing bias  
in more detailed  
classification

## The bias caused by the increase of the threshold - example of horticultural enterprises specialised to berry production

- The increase of the threshold from 2 000 euros to **10 000** euros cuts off 34 % of berry farms, 6% of production land BUT only 2% of the total crop.
- With special production such as with blueberry farming the cut off of the farming land is 35-38% and in crops 28 %
- In areas where the size of the horticultural farms are larger the impact of the cut-off is smaller.
- Berry farms are often small in relation to the value of production thus the cut off affects them more severely. They are also more concentrated to the Eastern Finland
- In some areas e.g. Kainuu (Northern-Eastern Finland) horticultural farms main production is in berry farming – the increase of the threshold does not have any impact on open land production as the farms that are left out are classified as berry farms although they may have some other small-scale production to the market

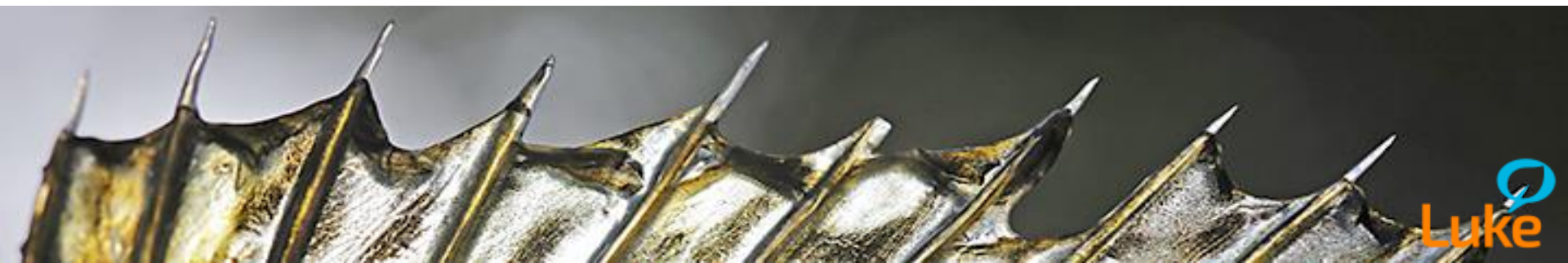
# Methods for bias correction under investigation

## Plausible methods:

- Donor imputation of the nearest neighbourhood
  - Can be too difficult with the rare production branches and with small farms
- Estimation methods based on register information on farms, growth region, local rainfall and weather statistics
- Mixed method of expert opinion based on donor imputation
- We are interested to hear about the experiences in official statistics and the recommendations of the statistical researchers

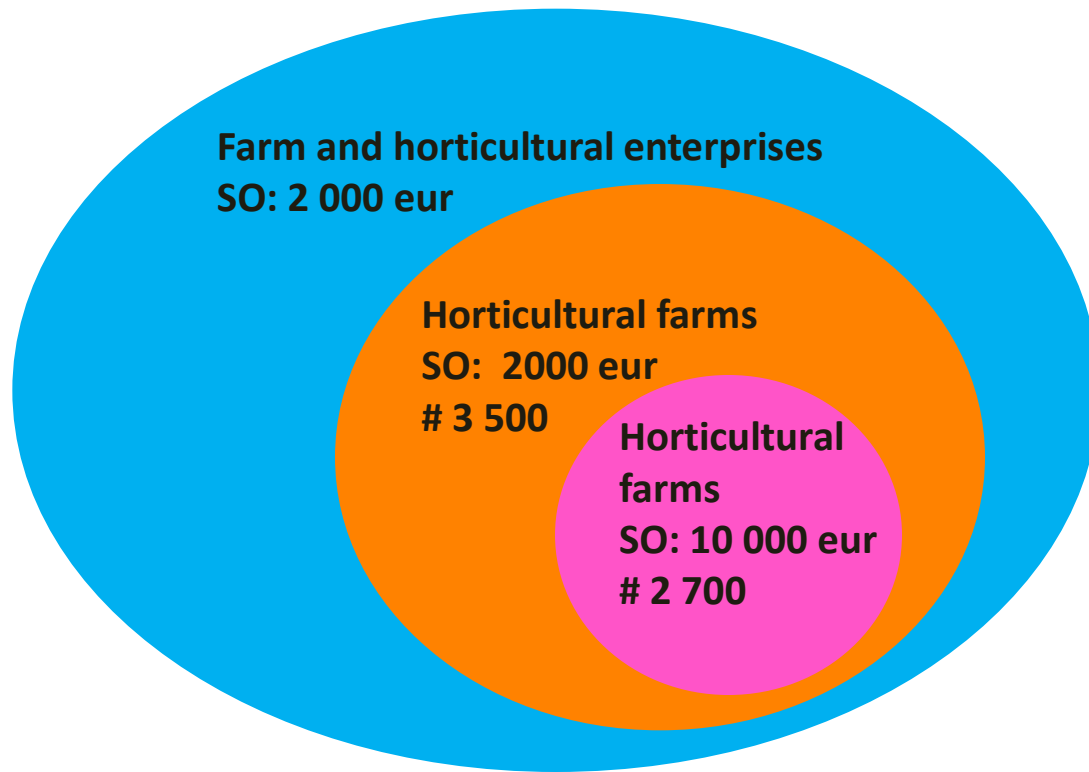
## Imputation methods out of focus:

- Imputation based on time series or the previous year
- Multiple imputation method has been excluded due to complexity
- Too complicated and resource intensive methods dependent on specific characteristics affecting the comparability of the statistics over the years





# Focus at present and in the future of horticultural production & change from details into main structural factors & economic value of production



*Improving efficiency by 800 units: savings of 4 000 eur - 8 000 eur & reduction of response burden by 23 %*






This is in our definition for a nursery



## Statistics service

The Natural Resources Institute Finland produces and publishes Finland's statistics on food and natural resources. These statistics provide information on the trends and current state of all fields in which the institute operates. The Natural Resources Institute Finland serves as one of Finland's four statistics authorities.



Agricultural

Forest

Fisheries and  
hunting

You will find us at <https://stat.luke.fi>

Thank you for your attention!  
- comments & suggestions ... ..

Contact:

[Forename.surname@luke.fi](mailto:Forename.surname@luke.fi)

Statistics:

<https://stat.luke.fi>