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):**
• 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%
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 intervies
- 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 up to 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 hectare 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

<table>
<thead>
<tr>
<th>Category</th>
<th>SO eur/ha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open field vegetables</td>
<td>11 399</td>
</tr>
<tr>
<td>Open field berries</td>
<td>4 874</td>
</tr>
<tr>
<td>Ornamentals in green houses</td>
<td>631 439</td>
</tr>
<tr>
<td>Vegetables in green house</td>
<td>647 148</td>
</tr>
<tr>
<td>Oat</td>
<td>499</td>
</tr>
<tr>
<td>Potatoes</td>
<td>4 421</td>
</tr>
</tbody>
</table>

*Threshold for EU-subsidies on green houses is 300 m²*
Sensitivity analysis:
Impact on estimated production volumes by SO

<table>
<thead>
<tr>
<th>Product</th>
<th>SO: 2 000 eur</th>
<th>SO: 6 000 eur</th>
<th>SO: 10 000 eur</th>
<th>SO: 15 000 eur</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># farms</td>
<td>ha</td>
<td>1 000 kg</td>
<td># farms</td>
</tr>
<tr>
<td>Strawberries</td>
<td>1 106</td>
<td>3 298</td>
<td>12 858</td>
<td>1 105</td>
</tr>
<tr>
<td>Highbush blueberry</td>
<td>147</td>
<td>72</td>
<td>115</td>
<td>117</td>
</tr>
<tr>
<td>All berries</td>
<td>1 612</td>
<td>5 797</td>
<td>15 575</td>
<td>1 251</td>
</tr>
<tr>
<td>Carrots</td>
<td>361</td>
<td>1 652</td>
<td>74 221</td>
<td>330</td>
</tr>
<tr>
<td>Garlic</td>
<td>67</td>
<td>19</td>
<td>49</td>
<td>52</td>
</tr>
<tr>
<td>All open land vegetables</td>
<td>1 408</td>
<td>9 142</td>
<td>185 284</td>
<td>1 251</td>
</tr>
<tr>
<td>Tomatoes</td>
<td>332</td>
<td>1 040</td>
<td>39 890</td>
<td>331</td>
</tr>
<tr>
<td>Butter-head lettuce</td>
<td>47</td>
<td>128</td>
<td>2 601</td>
<td>47</td>
</tr>
<tr>
<td>All greenhouse vegetables</td>
<td>566</td>
<td>2 250</td>
<td>83 058</td>
<td>565</td>
</tr>
<tr>
<td>Bedding plants (# 1 000 )</td>
<td>514</td>
<td>38 853</td>
<td>514</td>
<td>38 853</td>
</tr>
</tbody>
</table>
Sensitivity analysis: Impact of increasing the threshold from 2 000 eur to 10 000 euros, %

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

- 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 -&gt; imputation for adequate estimates

<table>
<thead>
<tr>
<th></th>
<th>SO: 2 000 eur</th>
<th>Diff: SO 2 000 eur/ 10 000 eur</th>
<th>% Diff 2 000 / 10 000 eur</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># farms</td>
<td>ha</td>
<td>1 000 kg</td>
</tr>
<tr>
<td>Strawberries</td>
<td>1 106</td>
<td>3 298</td>
<td>12 858</td>
</tr>
<tr>
<td>Highbush blueberry</td>
<td>147</td>
<td>72</td>
<td>115</td>
</tr>
<tr>
<td>All berries</td>
<td>1 612</td>
<td>5 797</td>
<td>15 575</td>
</tr>
<tr>
<td>Carrots</td>
<td>361</td>
<td>1 652</td>
<td>74 221</td>
</tr>
<tr>
<td>Garlic</td>
<td>67</td>
<td>19</td>
<td>49</td>
</tr>
<tr>
<td>All open land vegetables</td>
<td>1 408</td>
<td>9 142</td>
<td>185 284</td>
</tr>
<tr>
<td>Tomatoes</td>
<td>332</td>
<td>1 040</td>
<td>39 890</td>
</tr>
<tr>
<td>Butter-head lettuce</td>
<td>47</td>
<td>128</td>
<td>2 601</td>
</tr>
<tr>
<td>All greenhouse vegetables</td>
<td>566</td>
<td>2 250</td>
<td>83 058</td>
</tr>
<tr>
<td>Bedding plants (# 1 000 )</td>
<td>514</td>
<td>38 853</td>
<td>0</td>
</tr>
</tbody>
</table>

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
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

Farm and horticultural enterprises
SO: 2000 eur
# 3500

Horticultural farms
SO: 2000 eur
# 3500

Horticultural farms
SO: 10000 eur
# 2700

Improving efficiency by 800 units: savings of 4000 eur - 8000 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.

You will find us at [https://stat.luke.fi](https://stat.luke.fi)
Thank you for your attention!
- comments & suggestions … … …

Contact: 
Forename.surname@luke.fi

Statistics: 
https://stat.luke.fi