Towards automated use of satellite remote sensing data for crop production and agricultural land use statistics

The 106th DGINS conference on Earth observation for official statistics
Session 2: Applications of EO technologies and data in official statistics
Sub-session 2.1: Agriculture and Maritime economy

Maria Yli-Heikkilä

Using EO for agricultural statistics

1. Create reference data (field level, farm level)
2. Remote sensing data
3. Data-driven modelling (it works!)
4. Aggregation of results (needs of stakeholders)
5. Error analysis (in quest of sophisticated methods)

Uptake of EO for (agricultural) statistics takes time!

1. Create reference data (field level, farm level)
2. Remote sensing data 60%
3. Data-driven modelling (it works!) 10%
4. Aggregation of results (needs of stakeholders) 10%
5. Error analysis (in quest of sophisticated methods) 20%

EO for (agricultural) statistics **requires multi-domain knowledge!**

1. Create reference data
   - Domain + data science with GIS expertise
2. Remote sensing data
   - Remote sensing, IT
3. Data-driven modelling
   - Machine learning
4. Aggregation of results
   - Domain + data science with GIS expertise
5. Error analysis
   - Statistics

Using EO for agricultural statistics

1. Create reference data (field level, farm level)
   - Observational unit
   - The more the better
2. Remote sensing data
3. Data-driven modelling (it works!)
4. Aggregation of results (needs of stakeholders)
5. Error analysis (in quest of sophisticated methods)

Using EO for agricultural statistics

1. Create reference data (field level, farm level)
2. Remote sensing data
   - Huge volume of data
   - Cooperation in usage and storage
3. Data-driven modelling (it works!)
4. Aggregation of results (needs of stakeholders)
5. Error analysis (in quest of sophisticated methods)

Using EO for agricultural statistics

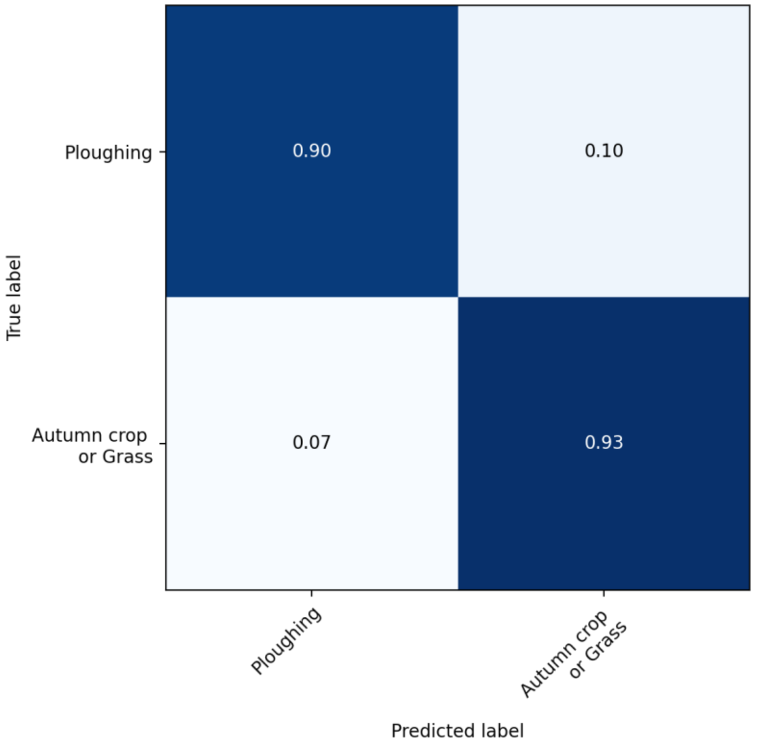
1. Create reference data (field level, farm level)
2. Remote sensing data
3. Data-driven modelling (it works!)
   - Machine learning
   - Requires high-performance computing
4. Aggregation of results (needs of stakeholders)
5. Error analysis (in quest of sophisticated methods)

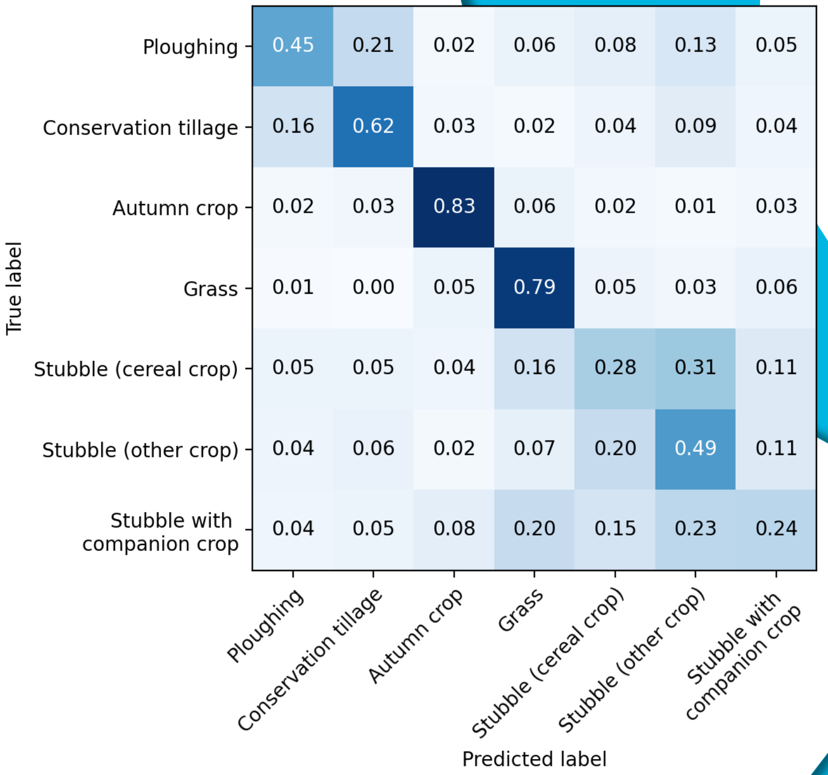
Example 3.a.: Crop yield statistics

Mean crop yields (kg/ha) over 2018–2020 in comparison with the mean post-harvest RF and LSTM predictions and the difference (delta) of the predictions from the crop statistics for the same period.

Crop type	Crop statistics	RF	RF Δ (%)	LSTM	LSTM Δ (%)
Winter wheat	4,186	4,240	54 (+1%)	4,078	-108 (-3%)
Spring wheat	3,446	3,547	100 (+3%)	3,836	390 (+11%)
Rye	3,660	3,660	0 (0%)	3,698	38 (+1%)
Feed barley	3,650	3,536	-113 (-3%)	4,125	475 (+13%)
Malting barley	3,713	3,603	-109 (-3%)	3,770	57 (+2%)
Oats	3,466	3,424	-41 (-1%)	3,701	234 (+7%)

Example 3.b.: Off-season soil cover statistics

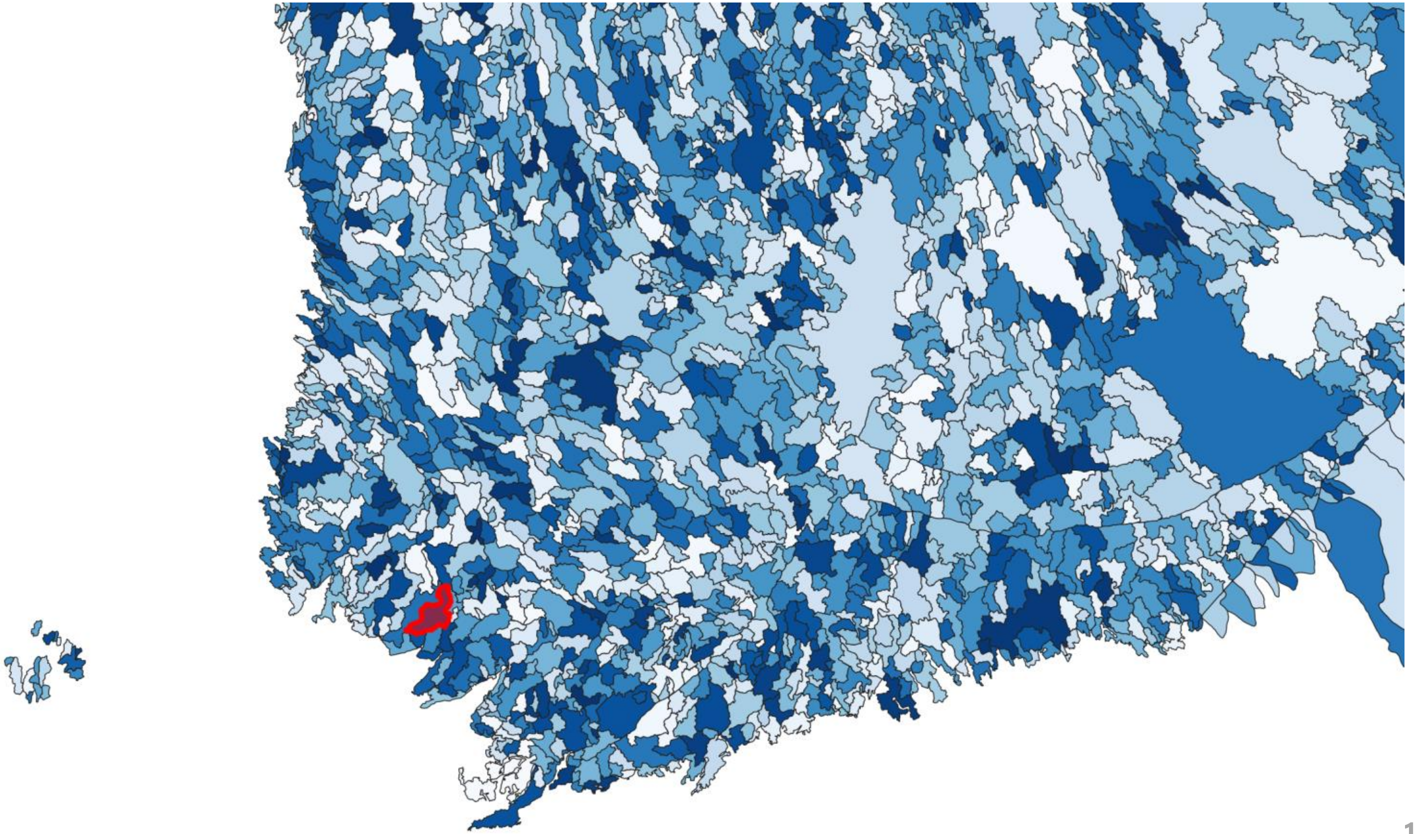




4. Aggregation of results

1. Create reference data (field level, farm level)
2. Remote sensing data
3. Data-driven modelling (it works!)
4. Aggregation of results (needs of stakeholders)
   – Preferrably dynamic spatial aggregation
5. Error analysis (in quest of sophisticated methods)

Example 4.a.: Off-season soil cover, case Savijoki



5. Error analysis

1. Create reference data (field level, farm level)
2. Remote sensing data
3. Data-driven modelling (it works!)
4. Aggregation of results (needs of stakeholders)
5. Error analysis (in quest of sophisticated methods)
   - Uncertainty properties of results important

Conclusions

- Uptake of EO for (agricultural) statistics takes time and is to certain extent scalable across other statistical domain (but not entirely)
- Automated pipeline requires knowledge of several disciplines
- Partly automated pipeline, constantly evolving

NATURAL RESOURCES
INSTITUTE FINLAND

Luke

NATURAL RESOURCES
INSTITUTE FINLAND