Policy Paper
From Yield Data to Yield Intelligence: Unlocking harvester data and digital twins for smarter agriculture

Authors: Isabelle Piccard, Hanna Huitu, Koen Uyttenhove, Sofie Landschoot, Sarah Verbeke, Wout Van Lierop

RILab: ScaleAgData RIL3 – Yield Monitoring

Date: July 2026

1 Highlights

The Issue: Europe is seeking to strengthen food-system resilience, improve evidence-based agricultural policy and accelerate the digital transformation of agriculture. Reliable and scalable yield intelligence is an important enabler of these objectives.

The Findings: The Yield Lab explored two routes: using harvester data with EO and weather data to estimate yield locally and regionally; and using digital twins to support crop-management decisions.

The Recommendations: Policies should move beyond data access alone and support traceable, quality-assured machinery data, interoperable sharing mechanisms and continued validation of yield-monitoring tools.

The Impact: Better calibrated and better documented yield data can improve precision farming services, regional yield indicators and evidence for policy evaluation.

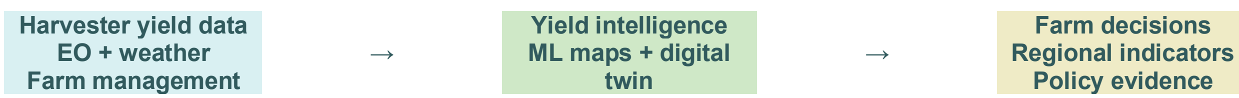

Figure 1. Yield Lab logic: turning farm-generated data into local services and regional indicators.

2 Context of the Issue

Agriculture is increasingly data-driven. Modern harvesters, Earth Observation systems and digital tools generate large volumes of information that could improve both farm management and agricultural monitoring. Yet this potential remains only partly used because data are scattered, insufficiently harmonised or difficult to interpret.

The ScaleAgData Yield Lab started from two research questions. First, can the large potential of harvester yield data be unlocked for yield estimation at both local and regional scale? At field level, the aim is to provide yield-variability maps, including for farmers who do not have a harvester equipped with a yield sensor. At regional level, the aim is to aggregate harvester-based and modelled field estimates into complementary yield indicators for crop monitoring and statistics.

Second, the Yield Lab explored whether digital twins can support precision-farming decisions. By combining management data, weather information and remote-sensing observations in a crop model, a digital twin can forecast yield potential and nitrogen uptake and support more targeted fertilisation decisions.

A central lesson is that data access alone is not enough. Machinery providers are increasingly open to sharing data with the consent of farmers as data owners, and harmonisation is progressing. However, the reliability of machinery-derived yield data remains a critical bottleneck. Users often lack information on how data were collected, whether the machine was calibrated correctly, which settings were used and which field conditions influenced the measurements. More automated quality recording and richer ancillary metadata are needed to enable filtering, correction and meaningful reuse.

This challenge is especially visible in potato production, where tare – soil harvested together with the potatoes – can distort yield maps. The Yield Lab therefore tested image-based tare estimation to improve net yield estimates and the trustworthiness of harvester-derived data.

3 Policy Recommendations

1. Establish common requirements for traceable machinery-derived yield data

Develop common metadata and quality requirements for harvester data, including calibration status, machine settings, sensor status, field conditions and data-processing steps. Where possible, these records should be created automatically to reduce dependence on manual input.

2. Strengthen fair and trusted agricultural data-sharing ecosystems

Support agricultural data spaces and interoperable APIs that allow machinery data, EO data and management data to be shared with the explicit consent of farmers while keeping data ownership and usage conditions clear.

3. Support integration into operational agricultural services

Promote yield maps and digital-twin outputs first as decision-support tools for farmers, advisors and service providers. Integration in FMIS, crop-monitoring platforms and advisory workflows will be essential for practical uptake.

4. Fund validation before operational policy use

Continue multi-year, multi-region validation of ML yield models, tare correction and digital twins before using outputs for official monitoring or policy assessment. Confidence metrics and uncertainty reporting should become standard outputs.

4 Evidence and Analysis

• Machine-learning yield estimation combines harvester yield data with EO and weather data to map potato yield and yield variability. The approach can support farmers in identifying high- and low-yielding zones and can be aggregated to regional yield indicators.

• The digital twin combines management data, weather and remote-sensing information to forecast winter wheat yield and nitrogen uptake, creating a pathway towards in-season fertilisation support.

• Image-based tare estimation addresses a key potato data-quality issue by estimating the soil harvested with potatoes. It can improve net yield maps, but more calibration and validation data are needed across soil types, moisture conditions and regions.

• The work remains at research and demonstration stage. Important limitations include noisy harvester data, incomplete information on crop variety and management, limited calibration datasets and the need to test model transferability beyond initial regions.

Data access	Harmonisation	Trustworthiness	Validation
Consent-based sharing	Common formats and APIs	Calibration, settings and metadata	Field testing and uncertainty

Figure 2. Four enabling conditions for trustworthy yield intelligence.

5 Strategic Alignment

The Yield Lab contributes to broader European objectives under the European Green Deal and the digital transformation of agriculture by improving the availability, quality and usability of yield intelligence for decision-making.

Policy area	Relevant link	Policy message
Agricultural policies	Vision for Agriculture and Food; current CAP; FSDN; crop monitoring; future CAP discussions	Yield intelligence can support productivity, resilience and resource-use monitoring, once validated.
Digital and data policies	Digital Decade, Data Union Strategy, agricultural data spaces	The main need is trusted, interoperable and quality-assured farm-data reuse.
Research and innovation	Horizon Europe, Agriculture of Data, smart farming and AI innovation	Further research is needed to move from promising prototypes to operational services.
Sustainability goals	Resource efficiency, soil protection, climate adaptation, SDGs 2, 9, 12, 13	Better yield and nitrogen information can support targeted management and more robust monitoring.

Initiatives under Horizon Europe, including the Agriculture of Data partnership, could provide a pathway for scaling and wider adoption.

6 References

• ScaleAgData RIL3 Demonstrator 3.1 – Improved potato tare estimator.

• ScaleAgData RIL3 Demonstrator 3.2 – Potato and winter wheat yield estimation.

• ScaleAgData architecture and data-governance documentation.

7 Contact Details

VITO Remote Sensing, Mol, Belgium

Isabelle Piccard

isabelle.piccard@vito.be

Consortium partners: VITO, UGent, AVR, CNH and LUKE.