Analysis of Case studies

D 5.2

December 2025

Funded by the European Union

Tools4CAP has received funding from the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement No. 101086311. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

ANALYSIS OF CASE STUDIES

Project name	Tools4CAP: Innovative Toolbox empowering effective CAP governance towards EU ambitions
Website	https://www.tools4cap.eu/
Document type	Deliverable
Status	Submitted
Dissemination level	Public
Authors	Emil Erjavec (UL), Ilona Rac (UL)
Contributors	Roel Jongeneel (WUR), Ana Gonzalez Martinez (WUR), Marcel van Asseldonk (WUR), Carla Wember (IFLS), Giorgos Charvalis (NP), Peter Witzke (EuroCARE), Juliana Arbelaez Gaviria (Chechglobe), Christian Schingo (ABACO), Trevor Donnellan (TEAGASC), Paula Palma Molina (TEAGASC), Emma Dillon (TEAGASC), Barbara Wieliczko (ERDN), Alberto Gutiérrez García (Itacyl), Zsolt Szabó (AKI)
Work Package Leader	University of Ljubljana
Project coordinator	Ecorys Belgium

This license allows users to distribute, remix, adapt, and build upon the material in any medium or format for noncommercial purposes only, and only so long as attribution is given to the creator.

Disclaimer: This document was produced under the terms and conditions of Grant Agreement No. 101086311 for the European Commission. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Table of contents

1. Introduction 5

1.1. Tools4CAP project introduction 5

1.2. The objectives and content of the deliverable 5

1.3. Methodology 6

1.3.1. General approach 6

1.3.2. Tune-up Evaluation framework 7

1.3.3. Categories of tools implemented within the Case studies 8

1.3.4. Criteria, sub-criteria and evaluation questions 9

2. Results of the evaluation 11

2.1. Quantitative tools 11

2.1.1. Summary of the Case study reports 11

2.1.2. TEF analysis 13

2.1.3. Key findings 22

2.2. Qualitative and participatory tools 23

2.2.1. Summary of the Case study reports 23

2.2.2. TEF analysis 25

2.2.3. Key findings 31

2.3. Monitoring tools 32

2.3.1. Summary of the Case study reports 32

2.3.2. TEF analysis 34

2.3.3. Key findings 39

2.4. Integrated cases 40

2.4.1. Summary of the Case study reports 40

2.4.2. TEF analysis 42

2.4.3. Key findings 47

3. Conclusions and next steps 49

3.1. Key findings by tool category 49

3.2. Final assessment 51

4. References 53

5. Appendices – Case Studies Factsheets 56

List of figures

Figure 1: Tune-up Evaluation Framework criteria. 10

List of tables

Table 1. Tune-up Evaluation Framework. 7
Table 2: Categorisation of tools implemented in case studies in the Tools4CAP project in 2025 8

Acronyms

AKIS	Agricultural knowledge and innovation system
AECM	Agri-Environment-Climate Measures
AGMEMOD	AGricultural MEmber State MODelling
CAP	Common Agricultural Policy
CAPRI	Common Agricultural Policy Regional Impact
CREA	Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria
CMEF	Common monitoring and evaluation framework
CS	Case studies
CSP	CAP Strategic Plan
CV	Cumulative Voting
DAFM	Department of Agriculture, Food and the Marine
DCE	Discrete Choice Experiment
DP	Direct Payments
EO	Earth Observation
EU	European Union
FADN	Farm Accountancy Data Network
FSDN	Farm Sustainability Data Network
FMIS	Farm Management Information Systems
FTE	Full-Time Equivalent
GAEC	Good agricultural and environmental conditions
GHG	Greenhouse gas
GIS	Geographic Information System
GLOBIOM	Global Biosphere Management Model
IIASA	International Institute for Applied Systems Analysis
IACS	Integrated Administration and Control System
IOI	Intervention-Objective-Impact
IL	Intervention logic

JRC	Joint Research Centre
LPIS	Land Parcel Identification System
MAGNET	Modular Applied GeNeral Equilibrium Tool
MS	Member State
ML	Machine Learning
NDM	New Delivery Model
NRPP	National and Regional Partnership Plans
NUTS	Nomenclature of territorial units for statistics
OECD	Organisation for Economic Co-operation and Development
PMEF	Performance Monitoring and Evaluation Framework
PMT	Policy monitoring tool
SDG	Sustainable Development Goals
SEA	Strategic Environmental Assessment
SMART	Specific, Measurable, Achievable, Relevant, and Time-Bound
SOC	Soil Organic Carbon
TEF	Tune-up Evaluation Framework
Tools4CAP	Innovative Toolbox empowering effective CAP governance towards EU ambitions
TPB	Theory of Planned Behavior
WUR	Wageningen University and Research

1. Introduction

1.1. Tools4CAP project introduction

In the context of a changing Common Agricultural Policy (CAP), the New Delivery Model (NDM), established under Regulation EU 2115/20211, introduces the new CAP Strategic Plan (CSP) framework and associated monitoring, review and evaluation requirements. The CAP (2023-27) has introduced a performance-oriented approach, to replace the compliance-based approach. Furthermore, additional responsibility and flexibility has been given to Member States (MSs) in the design of CAP Strategic Plans (CSPs) and the associated Performance Monitoring and Evaluation Framework (PMEF).

MS were obliged to acquire new capacities and set up appropriate methodological tools to ensure: i) alignment with the CAP, Green Deal objectives, SDGs and international commitments; ii) increased sustainability ambition compared to the previous programming period; iii) internal and external coherence of the intervention strategy; iv) consistency from regional to national levels; v) accountability through SMART targets and reflected in funding allocation; vi) reliance on ex-ante evaluations; vii) wide stakeholder engagement and coordination with regional bodies, socio-economic partners, and environmental and climate authorities; viii) process quality, simplification, and modernisation; and ix) cost-effectiveness of the adopted systems.

The Tools4CAP project provides CAP decision-makers with suitable tools for more evidence-based policy making, ultimately improving their capacity to design next generation SPs, and to improve the use of monitoring data. Although the European Commission has recently proposed the integration of CAP strategic planning into broader National and Regional Partnership Plans (NRPP, COM (2025/650), the use of tools to support strategic policy formulation, monitoring, and ex-ante evaluation remains relevant – perhaps even more so.

To foster the uptake of such tools, Tools4CAP aims to engage tools' end users to provide them with methodological guidance and capacity building opportunities. As such, this deliverable provides such a guide, building on the collective experience of the case studies (CS) conducted within the project's replication lab.

1.2. The objectives and content of the deliverable

The aim of this document is to report on the implementation of 12 case studies (Table 2) undertaken as part of the Tools4CAP project and provide a transversal assessment of their results. Based on set-up plans (D5.1), case studies were conducted to simulate the execution of a selection of tools that could potentially support the design (and implementation) processes of the CSPs. The case studies operationalized the following tools: (i) quantitative tools (mathematical simulation models and experimental economics), (ii) participatory decision tools for stakeholder-informed CSP design, and (iii) monitoring solutions to control compliance and support performance review and CSP ex-ante evaluation. In addition to cases testing individual tools, two integral examples included tools from all three of the above groups.

For each case study, an implementation report was produced, along with a descriptive case study factsheet, attached to this document. Deliverable 5.2 provides a comprehensive analysis of both the potential and the limitations of the tools that were tested, with a particular focus on the utility of the methods from an end-user (policy maker) perspective. The assessment is based on a modified benchmarking methodology previously developed under the project – the Tune-up Evaluation Framework (TEF) (D1.3).

The specific objective in this report is to synthesise the work undertaken, drawing on the experience and analysis of the CS to:

[1] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021R2115

- identify the possibilities for using tools for end users in terms of supporting strategic planning and harvesting monitoring data;
- define the possibilities and limitations of individual and integrated tools;
- highlight what needs to be considered by practitioners when using tools, including use of resources;
- identify possible upgrades to the tools;
- provide additional guidance for the update of methodological guidance (D2.3, D3.3 and D4.4).

1.3. Methodology

1.3.1. General approach

The analysis builds on the knowledge of completed project deliverables as of November 1, 2025 (selection and benchmarking of tools, methodological basis for individual tool groups, preparation and implementation of case studies). In addition, we draw on practical and theoretical insights associated with each tool or group of tools.

The scope of the evaluation is limited to the tools used in the CSs. These tools were identified in the inventory of tools developed in Task 1.1 of the project and were selected based on the interests of end users (following consultation with national project partners). Different types of tools were evaluated against a set of criteria and sub-criteria by triangulating information obtained from reports prepared by CS coordinators. Additional information and insights gathered from project partners during the project were also used. The CS results were presented in the form of an internal Tools4CAP project workshop, which, in addition to reviewing the reports, provided a broader understanding of the work undertaken and a validation of the CS process. Based on the CS reports, descriptive factsheets were also produced (attached to this deliverable), each one focusing on the tools used, outlining the new knowledge developed from the implementation of the tools and an appreciation of their usefulness. This evaluation is a synthesis of insights and findings stemming from the CS reports and these internal exchanges.

In line with the evaluation objectives presented above, the results of the evaluation focus on the potential and limitations of adopting different types of tools, as well as practical aspects that need to be taken into account when implementing such tools at the Member State (MS) level. The evaluation results and case study reports, therefore, can help end users in the MSs account for the various benefits and drawbacks, as well as good practices to be aware of, when adopting different tools. As such, the evaluation results can also help guide the choice of tools by end users.

The evaluation approach followed five steps. Steps 3 and 4 build on the Tune-up Evaluation Framework (TEF), which was developed for the assessment of tools in D1.3 and adapted to the needs of this evaluation. It serves as a conceptual and methodological instrument for the analysis of specific tools implemented in the case studies (the TEF is described in detail in Section 2.2).

In the first step, a suite of ***case study reports*** was produced to showcase experiences and outcomes that emerged during case study implementation. The reports outline the experiences with the implementation and key benefits, limitations and risks associated with the application of tools, as well as provide additional theoretical and methodological insights, where applicable.

In the second step, ***a grouping of case studies by category of tool implemented*** was undertaken. Four categories of tools had been pre-defined, reflecting the basic methodological and technical characteristics of the tools. The evaluation was performed for each group of tools separately, combining the assessment for tools with similar characteristics. This facilitated comparison among tools with similar characteristics.

In the third step, ***the evaluation framework*** based on D1.3 was defined, consisting of five criteria, further subdivided into 12 sub-criteria. To facilitate the evaluation according to the sub-criteria, one evaluation question was formulated for each sub-criterion. The answers to the evaluation questions under each criterion would then allow an overall judgment of the sub-category of tools against that specific criterion.

In the fourth step, we ***applied the evaluation framework*** at the level of each sub-category of tools by answering the evaluation questions, through the triangulation of the different information sources. Based on the answers to the evaluation questions, an overall qualitative judgment was formulated for each sub-category of tools.

In the fifth step, based on the evaluation results, ***conclusions and recommendations*** were elaborated for the use of tools in MSs and for the final work on the project (Handbook, D5.3).

1.3.2. Tune-up Evaluation framework

The evaluation was achieved by applying the Tune-up Evaluation Framework (TEF), which provides a conceptual and methodological instrument for the analysis of tools within the case studies implemented, specifically employed for the design and monitoring of the CSPs. For its operationalisation, the TEF relies on a categorisation of the tools, the definition of an evaluation matrix including criteria, sub-criteria and evaluation questions, and the triangulation of data from multiple sources.

Table 1 outlines the TEF and shows the different categories of tools evaluated, as well as the criteria, sub-criteria and related evaluation questions used to guide the evaluation. Thus, for each category of tools, 12 pre-defined evaluation questions were answered, with each evaluation question belongs to a specific sub-criterion. The overall judgment of an evaluation criterion was formulated based on the answers to the evaluation questions belonging to that criterion (e.g., the judgement of Accuracy is based on the answers to EQ 1 and EQ2).

It is important to clarify that, while the definition of the criteria and sub-criteria was the same for all categories of tools evaluated, the evaluation questions differ slightly among categories of tools (the same approach as in D1.3). This is due to the different practical and technical implications among very different tools, which required the questions to be tailored to the specific methodological and technical characteristics of each category of tools. For example, granularity and accuracy are characterised differently when discussing quantitative modelling tools and participatory tools.

However, in all cases the evaluation questions are devised to assess the criteria, whose underlying definitions remain the same across tools. Below, we describe the grouping of CS by category of tool (Section 1.3.3) and the definition of the evaluation matrix, whereas the data collection activities are described in the next section.

Table 1. Tune-up Evaluation Framework.

Criteria	Sub-criteria	Evaluation Questions (EQ)
Accuracy	*Data accuracy and validity*	EQ1
	Replicability	EQ2
Reliability	*Replicability*	EQ3
	Transparency	EQ4
Applicability	*Multifunctionality*	EQ5
	Experience	EQ6
	Granularity	EQ7
Accessibility	*Ease of use*	EQ8
	Tool availability	EQ9
Efficiency	*Time efficiency*	EQ10
	Cost effectiveness	EQ11
	Human resources efficiency	EQ12

Source: Tools4CAP 2025

1.3.3. Categories of tools implemented within the Case studies

In Deliverable D.5.1, we described the individual countries' choice of tools and implementation plans. The factsheets attached in the appendix briefly report on how the case studies were conducted. Table 2 presents the selection of tools by case study and their classification into tool categories. The integral cases carried out in the Netherlands and Slovenia are so specific that we treat them as a separate group although they include elements of other groups of tools.

Table 2: Categorisation of tools implemented in case studies in the Tools4CAP project in 2025

Case study	Tools used	Quantitative tools	Qualitative tools	Monitoring tools	Integral case study
Czechia	GLOBIOM model	X			
France	Experimental economics	X			
Germany	IOI Matrix		X		
Greece	Policy monitoring tool			X	
Hungary	CAPRI model	X			
Ireland	Intervention logic		X		
Italy	Policy monitoring tool			X	
Netherlands	IOI Matrix, Joint Eco-scheme-farm income model, Policy monitoring tool	X	X	X	X
Poland	Prioritization tool for needs assessment		X		
Slovenia	Vision building, Policy indicators, Scenario analysis,	X	X	X	X
Spain CS 1	CAPRI model	X			
Spain CS 2	Soil organic carbon map index			X	

Source: Tools4CAP 2025

As described in Deliverable D 5.1, the selection of tools for case studies was made from the set of tools defined in Deliverable D 1.3 based on the declared interests of end users within the CS countries. Case study coordinators identified the interest of potential end users through direct consultation with them (mainly agricultural ministries, but also other government departments and stakeholders involved in the agricultural policy process).

Tools implemented in the case studies were grouped into three main categories, which reflect the basic division of tools set out with the project logic. The specific tools or subtypes of tools applied are listed in parentheses:

1. ***Quantitative tools*** (mathematical modelling tools, experimental economics);
2. ***Qualitative participatory tools*** (participatory vision building, Intervention logic, prioritisation of needs, scenario building);
3. ***Monitoring tools*** (different policy indicators);
4. ***Integral case studies*** (a special category, which includes and combines topic-focused tools from the three basic groups of tools).

Quantitative tools: These tools serve to generate quantitative evidence through the analysis of policies, either ex-ante or ex-post, to inform decision-making. For the purposes of this analysis, this category is further divided into two sub-categories of tools: experimental economics and mathematical modelling tools. Experimental economics is primarily used as a tool for testing the preferences of participants before agricultural policy measures are introduced. It examines farmers' responses to measures, whether they might lead to the desired behaviour and uptake, and assists in the planning of measures. Mathematical modelling tools primarily focus on simulating scenarios, interventions, or impacts. They can be used to predict the outcomes of various policy decisions before they are implemented. In this assessment, we discuss the CAPRI model and GLOBIOM (Global Biosphere Management Model) in the chapter on monitoring tools, and the FARMDYN and AGMEMOD models in the chapter on integral CS.

Qualitative participatory tools are based on qualitative methodologies that enable the identification and assessment of stakeholder perspectives and needs, which can be used to inform policy assessments and policy choices. The tools are very diverse and may also be supported by quantitative techniques.

- Voting and prioritisation tools refer to a spectrum of instruments and methodologies, such as number voting, cumulative voting, or hierarchical prioritisation, that employ voting and/or analysis to arrange or prioritise tasks. They are used in decision-making processes to prioritise or rank items, options, or needs by assigning them numerical values.
- Participatory vision building supports the joint development of a vision for strategic policy planning with key stakeholders. The method helps to create a shared vision of the future that everyone understands and supports, through a process of interactive workshops and the preparation and adoption of strategic documents.
- Intervention logic shows how specific measures contribute to the achievement of strategic objectives (e.g., increased productivity, environmental protection). It is a structured approach to agricultural policy planning that links objectives, interventions, and results.
- The Interventions-Objectives-Impact (IOI) matrix is a table that helps to understand and determine the links between interventions and the achievement of agricultural policy objectives. It is informed by desk research, complemented by expert opinion, where other information is lacking.
- Scenario building is a useful method in agricultural policy that facilitates discussion of the possible outcomes of different support packages for achieving objectives.

Monitoring tools, in this classification, serve to collect and make available the necessary data for the performance review of the CSPs (performance monitoring tools). They may rely on different methodological and technical or technological approaches, and they play an integral role in achieving result indicators, by providing evidence on whether Member States are progressing toward, or remain on track to meet, the targets set out in their CSP. These tools are used in line with the Commission's prescribed methodologies setting out the use of specific data to calculate the result and impact indicators[2].

Integral case studies combine different tools to address a specific area (e.g., Dutch eco-schemes) or inform a specific element of strategic planning (Vision building for Slovenia).

1.3.4. Criteria, sub-criteria and evaluation questions

This section describes the proposed criteria, sub-criteria, and the related evaluation questions of the TEF methodology applied in Deliverable D1.3. As shown in Figure 1, a total of five criteria were defined to structure the analysis and ensure its comprehensiveness. The selection was made to draw operational conclusions regarding the technical and methodological aspects that can enable or hinder the practical implementation of the tool. The set of criteria provides a strong basis for identifying adequate evaluation questions aimed at pinpointing the characteristics or aspects of tools that contribute to their successful application. The criteria, therefore, aim to outline the potential and limitations of the tools, which in practice reflect strengths and weaknesses to be accounted for when selecting and implementing them.

The five criteria are defined below.

Accuracy. This reflects how successful the tool is in offering results that reflect reality, without any distortion or omission of those results. Results that are distorted or omit critical information can lead to misinformed decisions based on flawed assumptions, which can have serious consequences (e.g. waste of time and resources, missed opportunities, increased exposure to risks, stakeholder discontent, damages to reputation, failure of public policies).

Reliability. This measures the degree of consistency and stability of the results obtained when implementing the tool. This criterion measures the extent to which the process of implementing a tool produces consistent findings when repeated. A tool that generates inconsistent results through repeated use can be affected by non-systematic bias that hinders the validity and generalisability of the results. Conversely, systematic bias may produce consistent but distorted results, undermining accuracy even when reliability appears high.

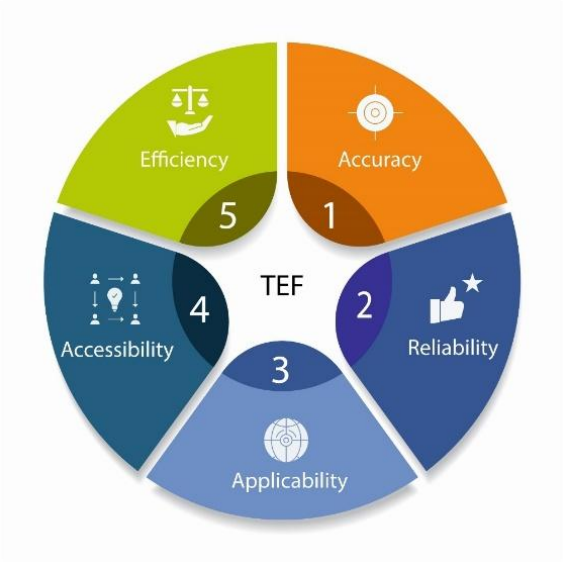
[2] https://agriculture.ec.europa.eu/system/files/2023-09/pmef-result-indicators_en.pdf

Applicability. This assesses the extent to which the tool can be used in different contexts and pursue different objectives. It shows the potential that a tool has to achieve different goals and perform different tasks. Gaining insights into the application of the examined tools helps end users make informed choices about which tools to adopt in a given context.

Accessibility. This shows the extent to which the tool and its underlying methodologies, data, assumptions, and outputs are clear and accessible to end users. This criterion is important to provide an understanding of how feasible it would be to transfer the use of certain tools to other contexts. The evaluation explores limitations that could hinder the adoption of certain tools by end users.

Efficiency. This measures the extent to which the tool can achieve desired outcomes given the resources available, including time, human resources, and costs. Moreover, this criterion is used to understand whether some tools achieved similar (or greater) benefits at the same/lower cost in terms of resources allocated for the tool's set-up, maintenance and implementation.

Figure 1: Tune-up Evaluation Framework criteria.



2. Results of the evaluation

This chapter reports the results of the evaluation for each category of tools used in the case studies. For each category of tools, we provide a short description of the tools and the policy tasks for which the tools were used. Then, we report a summary of results for each sub-category of tools according to the five evaluation criteria, and the full analysis based on the answers to the 12 evaluation questions (organised according to the five criteria and 12 sub-criteria). Lastly, for each category of tools, a section of the chapter brings together the key findings from the corresponding sub-categories of tools and provide an overall judgement of the category of tools against the five evaluation criteria.

2.1. Quantitative tools

2.1.1. Summary of the Case study reports

Case study Czechia – Impact assessment of various CAP policy instruments using the GLOBIOM model for the Czechia

The Czechia must align EU objectives (sustainability, climate resilience) with national ambitions for its agricultural sector and its agricultural competitiveness. The CAP Strategic Plan (2023–2027) allocates 54% of funds to environmental and climate objectives. The Czech CS (appendix) aims to assess the impact of the CAP Strategic Plan (2023–2027) using the GLOBIOM model for Czechia. The model analyses the economic and environmental effects of policies and their synergies and trade-offs in the context of climate change. Therefore, scenarios have been developed for different policy options (no CAP, DP (direct payments) only, GAEC (good agricultural and environmental condition) conditions only, comprehensive CAP (DP+GAEC)) under different levels of warming in the context of agricultural development up to 2050. The impacts of CAP instruments have been quantified in terms of competitiveness, sustainability, and resilience.

The fundamental methodological work involved the integration of specific data (direct CAP payments under Pillar 1, GAEC production conditions, and climate impacts) into the GLOBIOM model for Czechia. The results show the importance of environmental conditions (GAEC) for some important environmental indicators (share of permanent grassland, reduction of greenhouse gas emissions). Direct payments increase emissions from certain crops (e.g., potatoes, wheat). The economic effects indicate changes in the production structure (by 2050, potatoes and oilseed rape will gain competitiveness, while barley will lose competitiveness). The recommendations particularly emphasise the importance of environmental conditions for payments.

The case study suggests that some key decision-makers (the Ministry of the Agriculture) have limited interest in this type of analysis, while on the other hand, interest was shown by the Ministry of the Environment. The time and expertise required to implement the model were greatly underestimated by potential end users of the results. Resource requirements for the modelling are significant. The model can be used for CAP decision-making only if a permanent partnership is established between IIASA or Czechglobe (developers of the model) and the relevant Czech ministry, with intensive cooperation from a Czech research partner who would work on the model on a permanent basis and incorporate national specific factors in the model and facilitate its ongoing maintenance and development.

Case study Hungary – Use of the CAPRI model to assess the impact of CAP reform policy options on soybean production in Hungary

Hungary seeks to increase local production of protein crops (especially soybeans), which would contribute to a reduction of the EU's dependence on protein feed imports and assist in achieving sustainability goals. The new EU legislative proposal (COM 2025/650) indicates that the upper limit for coupled payments could be raised from 15% to 25% of the national budget for direct payments, which provides scope for greater incentives to be provided for soybean production. Therefore, a study was conducted using the CAPRI model to analyse the effects of increasing coupled payments for soybeans to determine the implications for land use and crop production, farmer incomes, and environmental indicators (carbon dioxide, methane and nitrous oxide emissions). The case study also tested the ad-hoc application of the CAPRI model, which is mainly used for EU ex-ante impact assessments of policies at the MS level. The model was appropriately adapted for scenario settings.

The key findings from the many scenarios undertaken (combining different support options for protein crops) are that scenarios with higher payments or more hectare restrictions significantly increase soybean production (up to +160%) and farm income (+80%), with a slight decrease in fodder crop production. At the same time, carbon dioxide emissions are reduced by up to -23% in scenarios with eco-schemes, and methane and nitrous oxide also decrease. Increased supply of soybeans leads to lowers soybean prices, but the impact on the EU market remains small, with the EU remaining a net importer of soybeans. The study recommends that the future CAP consider stable and targeted support for protein crops – increased payments and flexible hectare restrictions.

This case study also confirmed that the CAPRI model can be used as a standard tool for ex-ante analysis of CAP policies in national strategic plans. However, the case study found that interest among Hungarian state institutions in impact assessments and quantitative support is limited. The case study confirms that there is institutional support for this type of scenario analysis in principle, but there is less interest of the part of these institutions in concrete involvement in the formulation of questions and the assessment of results. There is also a clear dependence on the support of model developers when it comes to more complex EU-set models (in this case CAPRI) and the need for intensive and continuous investment by national research institutions in modelling, which is not possible without stable state support.

Case study Spain (1) – Use of Digital Farm Book data in the CAPRI model to assess pesticide policies in Spain

EU Directive 2009/128/EC requires the sustainable use of pesticides, the reduction of risks to health and the environment, and the promotion of integrated plant protection. The Spanish CSP includes measures (eco-schemes, statutory management requirement related to plant protection – to achieve these objectives. In Spain, data on pesticide use are collected through the Digital Farm Book (Farm management information system – FMIS), but reporting is voluntary, which limits the representativeness of the data collected in the context of the farm population. The aim of the case study was to verify whether data from the FMIS on pesticide use can improve the quality of the CAPRI model for ex-ante analysis of CAP policies, particularly in assessing production costs, the impact on crop structure, and environmental indicators (e.g., use of highly hazardous pesticides – HHP). For this purpose, the model was appropriately adapted and validated.

The key finding from the Spanish CS is that using FMIS data results in more realistic pesticide use patterns. FMIS data show low use of fungicides and high use of herbicides (glyphosate), reflecting the dry climatic conditions and the eco-scheme payments for conservation tillage in Spain. However, the limited coverage of FMIS data poses a challenge that also affects the quality of the model's input sources. The key risk is related to voluntary reporting for farmers, which results in heterogeneous data coverage across regions and crops and, hence, limited representativeness.

Nevertheless, the use of FMIS data does improve the robustness of modelling results and the policy relevance of scenarios. The transition to mandatory Digital Farm Book reporting could become the basis for better evidence-based agricultural policy in this context, if data reliability can be ensured and if the burden on farmers of providing this data is reduced. Therefore, it is recommended that policymakers introduce mandatory and harmonised Digital Farm Book reporting, with minimal administrative burden on the farmer. It is also necessary to establish procedures for data anonymisation, validation, and its integration into models. The key is to build trust among farmers by simplifying reporting, transparency in data use, and avoiding a punitive approach to policy compliance.

The study highlights the importance of diversifying data sources for complex models such as CAPRI, which can increase their value at national and EU levels. It also emphasises the importance of involving decision-makers in the administrative definition of new data sources that can improve the evidence-based approach in agricultural policy.

Case study France – Use of experimental tools to improve AECM in the Occitanie region

The Occitanie region in France manages CAP measures, including the Agri-Environment-Climate Measure (AECM) for the transition to greater protein autonomy (animal feed self-sufficiency) on livestock farms. Farmer participation in this measure is low (273 farms in 2025), so the aim of the study was to understand the reasons for low participation and the possibilities for adapting the measure. Therefore, in the French CS, experimental methods in the form of Discrete Choice Experiment (DCE) and subsequently the Theory of Planned Behavior (TPB) were tested to assess farmers' preferences of implementation modalities. The analysis aimed to provide recommendations for improving the design of the measure to increase its attractiveness and effectiveness.

DCE was initially selected to measure preferences regarding the attributes of the measure (payment terms, technical support, choice of practices). Due to limitations (at the request of regional administration representatives, it was not possible to include the monetary attribute, while the small number of farms involved made the DCE design unfeasible), the method was replaced by TPB and preference ranking. The survey questionnaire was sent to 273 farmers; 43 responses were received (16% response rate). The key findings were that farmers appreciated the environmental and production benefits of the measure, but financial and time constraints remained a major obstacle. The most desirable practices were the expansion of fodder legumes and the optimisation of grazing. When asked about technical support for the implementation of measures, the farmers surveyed emphasised that farmer autonomy or voluntary assistance would be preferable, while mandatory external support was less popular. In any case, farmers recommend simplifying administrative procedures and reducing payment delays. It was found that communication needs to be strengthened: promoting the benefits of protein autonomy, examples of good practice, and the exchange of experiences between farmers.

The limitations of the CS are the small sample size, the absence of a monetary attribute, and the limited influence of the region on the design of the measure. It appears that the use of experimental tools can be effective for ex-ante impact assessments and in the design of measures, but their use in the real world is quite constrained. Sufficient resources and stakeholder involvement are needed from the outset. Experimental economics methods such as DCE are demanding to implement (in terms of time, resources, and expertise) and work best in the policy design phase, rather than retroactively. In contexts with limited resources, simpler methods (focus groups, interviews) or surveys based on TPB may be more effective for gaining useful insights.

2.1.2. TEF analysis

Analysis by criteria

ACCURACY

Accuracy analysis considers how successful the tools are in offering results that reflect reality without any distortion or omission in those results. The results of the evaluation are presented below.

DATA ACCURACY AND VALIDITY

Data accuracy and validity refer to the extent to which data input and its sources are traceable and validated.

Evaluation Question: To what extent do the tools rely on well-founded evidence (e.g. that the sources of data used during the implementation of the tool are known to be reliable)? Is it possible to easily trace the information used to support the policy choices back to its original source? Are measures in place to avoid mistakes and redundancy and to ensure consistency in the data or information used to support the policy choices (e.g. to verify the quality and reliability of the information)? Are results triangulated, corroborated, or substantiated by other data sources?

Any modelling, whether complex simulation models (CAPRI and GLOBIOM) used here, or in the form of various experimental economics techniques, requires the use of a demanding set of data through which we attempt to understand and simulate real-life situations. The quality of any quantitative tool is determined by the quality of the data used, so special attention must be paid to data provision and data quality.

In the case of models presented here, operating at the macro level, thus covering multiple MSs, such as CAPRI[3] and GLOBIOM[4], the data sources used during the implementation of the tool are known to be reliable. CAPRI constructs its own database (COCO – complete and consistent) at the global, national, and regional levels. The databases exploit, wherever possible, well-documented, official, and harmonised data sources, especially data from EUROSTAT, UNFCCC, FAOSTAT, national data, and extractions from the FADN. The CAPRI model strives for consistency by using regularly updated data sources and constructing its database to ensure data consistency across different dimensions, such as market balances and spatial levels, matching physical and monetary data. The same applies to GLOBIOM, which uses complex and validated databases. Additionally, GLOBIOM regularly assesses the baseline and databases for EU MS using the CAPRI baseline as its reference.

[3] https://www.capri-model.org/dokuwiki_help/doku.php?id=the_capri_data_base

[4] https://iiasa.ac.at/models-tools-data/globiom

Despite the relatively good data infrastructure of the models, there are quite a few limitations. CSs have shown that for the purpose of specific analysis of strategic agricultural policy issues, it was necessary to further verify and supplement the models with data, in the context of agricultural production, input use and policy (CS Czechia, CS Hungary, CS Spain). This requires expertise and time to supplement the models to ensure that they are relevant. Large scale models usually demand more data than is available and may be processed with existing capacities. Furthermore, the models can be improved using newly available data sources (CS Spain 1) after thorough assessment of the data's completeness and representativeness at regional levels. In any case, investment in data provision is crucial for the quality and thus the usefulness of the modelling result relating to new agricultural policy issues, which are becoming increasingly differentiated. Large-scale models have reasonable data to assess markets, nutrient flows, GHGs, and health, but struggle to also cover emerging policy concerns, such as biodiversity, pesticides, antibiotics, and animal welfare, which require a huge amount of very granular data, ideally differentiated across space and time.

Conducting experiments in experimental economics entails a range of data acquisition activities, from focus groups and in-depth interviews used to develop questionnaires to the collection of survey data using these instruments. In strategic planning, there is rarely sufficient time to conduct such complex analyses, while researchers often find it difficult to compromise on methodological rigor.

Complex quantitative tools are often a challenge for strategic planning, as they require extensive preparatory work, while their applicability can vary depending on the question to be examined, as well as on the quality, availability, and suitability of data. Decision-makers often demand quick but reliable results; this requires intensive data collection, analysis, and validation done *in advance* of solving specific policy problems, which in turn means intensive public investment in data and research.

Tool Representativeness

The capacity of a tool to accurately and comprehensively capture or depict a system, dataset, or phenomenon. It combines the notions of accuracy and coverage in its depiction.

Evaluation Question: To what extent do these tools exhibit a high degree of sensitivity (i.e. how much the results of the tools are influenced by changes in the input data)? Are the tools able to capture the diverse facets, variations, or elements within the system or dataset they seek to represent?

The extent to which quantitative tools (whether mathematical models or experimental economics techniques) can actually represent the real and complex situations in agriculture and rural areas, covering the various dimensions and issues of agricultural policy, is a multi-dimensional issue. It depends on the data, the precision of model design, and on the model's implementation and interpretation. Models are only as valuable and useful as the effort and knowledge that their developers can invest in them and the extent to which communication with clients, especially users from the state administration, actually takes place. Therefore, continuous dialogue with clients and the improvement of models and the ways in which they are used is necessary. A representative modelling tool takes several years to develop and requires top-level knowledge, a lot of work, and dialogue with clients. The models used by the European Commission to support decision-making in the CAP (Barreiro-Hurle et al. 2024) are a good example of what is required. At the MS level, the use of specifically developed national models is less widespread and limited to certain countries such as The Netherlands, Ireland, Belgium, Germany, France, Slovenia (Tools4CAP, 2023a).

Large and complex models such as GLOBIOM, CAPRI, MAGNET, AGMEMOD, and others, have the potential of provide great granularity, as model structures are usually designed with or at least open to the inclusion of region- and crop-specific parameters to allow better tailoring to MS' circumstances. But in practice the resources are often insufficient to fully exploit this potential for tailoring, such that general rules/parameters are used that are uniform across regions or even across the EU.

If used at the MS level, the large-scale models require additional input in the preparation and validation of data and the models themselves (in terms of their sensitivity and ability to generate useful results for decision-makers). As these are complex tools whose primary purpose is to assess the impact of policy options at EU level, they are also difficult to adapt to the more specific needs of individual MSs. Above all, they generally require sophisticated handling and the ability to modify the tools, which calls for cutting-edge expertise and long-term investment in individual researchers or even modelling teams. Therefore, the continued maintenance of such tools is highly dependent on the development of research groups in individual countries and long-term investment in this type of work and scientific competence.

All three examples (Czechia, Hungary, and Spain) demonstrate that mathematical models show considerable credibility in their application at the MS level, which shows that the data and structure of the models are appropriately applied and calibrated to reflect the situation in the three MS in question.

The use of experimental economics methods, if carried out in accordance with scientific standards for these tools, can provide a good illustration of the current dispositions and potential behaviour of beneficiaries when designing agricultural policy measures. However, as the French CS has shown, it is difficult to secure genuine interest and resources for this type of research in a timely fashion, which is a prerequisite for it to be directly used in decision-making.

RELIABILITY

The assessment of reliability measures the consistency and stability of the results obtained with the application of the selected tools. This criterion measures the extent to which the process of implementation of the tools produces consistent findings when repeated. The results of the evaluation are presented below.

REPLICABILITY

Replicability refers to the extent to which the tool's results are (or can be expected to be) consistent across studies aimed at answering the same question, either by replicating the same approach or using new data or methods.

Evaluation Question: Can the results of the category of tools be replicated successfully by others or through different approaches? Do the results remain consistent with other studies and approaches?

Mathematical models, especially complex ones used in case studies, are so specific, based on different theoretical foundations and using different data sources, that it is difficult to ensure complete consistency of results, for example between models. Here, the focus is more on whether the directions of the results are similar and, depending on the nature of the models, their responsiveness/sensibilities. These comparisons have been undertaken in past (H2020 project "SUPREMA") and ongoing research projects (HE projects, "Brightspace", "Lamasus", "Act4Cap27"), usually with mixed results showing that the analysis of divergence in results between different models is time consuming, but also insightful. Another important criterion is whether the model also produces the expected results, based on theoretical assumptions, past modelling experience, and scientific sources. From this perspective, all three mathematical modelling CS have proven the models to be quite reliable and confirm the hypothesis that both large complex models (CAPRI and GLOBIOM) are suitable is assessing effects at the national level. The same conclusion likely applies to other models of this type. Unfortunately, we did not test the possibility of developing a farm-level model in a more complex or simplified version, but practice shows that these models can also be used effectively in CAP strategic planning (Tools4 CAP 2023b and 2024a).

Reproducibility of results is also a criterion for statistical analyses, including those employed in experimental economics; it is is a fundamental requirement of scientific work. However, the French CS did not quite reach the standard, as the sample was too small. In the end, only simple statistical analysis and qualitative analysis of open-ended questions were possible rather than the planned quantitative analysis of farmer preferences. Nevertheless, the results produced so far are credible and justify expanding the experiment to a larger sample. Here we encounter the dilemma highlighted several times in this report, namely, whether it is even possible to directly apply these types of tools for strategic planning without long-term planning and direct state investment, given the resource and time constraints typically associated with policy-making.

TRANSPARENCY

Transparency refers to the extent to which information concerning underlying methodologies, data, and results of the tool are available and accessible, as well as the extent to which the tool is open to external review.

Evaluation Question: is the information concerning underlying methodologies, data, and results of the tool available and accessible, and is the tool accessible for external review?

In large and complex mathematical models, such as CAPRI and GLOBIOM, it is considered good modelling practice to conduct reviews and testing to ensure their quality. Furthermore, the robustness of model results is typically verified through a supervisory committee or guidance group within the context of a project, as well as through an internal quality review conducted by the research institute performing the analysis. CAPRI and GLOBIOM typically come with documentation accessible to the public. The information regarding the underlying methodologies and data of macro-models like CAPRI and GLOBIOM is readily available and accessible, and the core version of these tools

is open for external review. Similarly, all linkages with well-established and harmonised databases are also well documented, increasing the transparency of the tools. While this requires a certain level of expertise in modelling, training materials are provided on the website, and training sessions are available for those who register. Although the input parameters (scenario descriptions) and results of some studies might be classified as sensitive and, therefore, not available for external review, the stable release version of CAPRI is offered for free download even without registration.

However, understanding the structure and functioning of large and complex models requires considerable knowledge and experience. Even skilled modellers need some time to familiarise themselves with how to implement a model scenario or interpret model results. This is especially the case when, for example, they take over the development of another model they have not constructed themselves. Therefore, complete transparency is extremely difficult to be achieved with such tools, and relies strongly on the trustworthiness of the research groups that develop and apply the model. It is also important to understand the theoretical foundations and programming environment in order to evaluate the realism of the results. New or potential users often do not trust the abstract depictions of reality that are offered by models and which are by necessity a simplification of reality, however complex. This gap in understanding between modellers and policymakers needs to be considered in the communication of results.

In experimental economics, transparency is generally ensured through the proper application of statistical techniques. As these approaches are currently mainly limited to scientific research when it comes to agricultural policy, it is important to point out that the methodology itself is transparent and often rigorously examined. These analyses undergo peer review, where some journals even require data sources and detailed descriptions of data processing procedures prior to publication, so that the procedures and results can be fully reproduced. The transparency of the methodology, data, and results is therefore ensured, and a research process undertaken in accordance with all recognised scientific standards can ensure accuracy and reliability. In the rather messier process of hands policymaking, this rigour poses a challenge, as users demand quick answers within an acceptable timeframe. The CS in France, where initial ambition had to be curbed and finally applied only simple statistical analysis, is of course not questionable in terms of transparency.

APPLICABILITY

The applicability assessment aims at measuring the extent to which the tool can be used in different contexts and to pursue different objectives. The results of the evaluation are presented below.

MULTI-FUNCTIONALITY

Multi-functionality refers to the extent to which the tool pursues several objectives and policy-making steps at the same time, (i.e., the number of policy objectives or policy-making steps that can be addressed by the tool), as well as the extent to which the underlying assumptions of the tool can be easily changed to address other policymaking steps or other objectives.

Evaluation Question: To what extent can these tools be used to address different policy objectives and steps of the policymaking process? Has the category of tools been used to contribute to one or more of the CAP Specific objectives and to one or more steps of the CAP Strategic Plan design or monitoring? Can the tool be tailored or re-adjusted to assess or contribute to the assessment of the other Specific Objectives and policymaking steps?

Both mathematical models tested in the context of the CSs can accommodate multiple objectives and policy instruments. CAPRI can also provide insights into potential interaction effects at the farm level resulting from the simultaneous implementation of different policy instruments. CAPRI and GLOBIOM provide information on various types of impacts, such as economic impacts (revenues, trade, production) and impact on environmental indicators (e.g., carbon dioxide, nutrient balances, ammonia emissions, biodiversity impacts).

The CAPRI model is well suited for evaluating the impact of scenarios relating to the CAP, trade, and environmental policies on agricultural production, income, markets, trade, and the environment, on both a global and regional (NUTS2) scale. It is frequently used to evaluate changes to the CAP and the potential impact of free trade agreements on the agricultural sector, as well as the impacts of other sectoral policies, such as those related to the

environment and climate change[5], and it can be adjusted to assess and contribute to the assessment of specific objectives in foresight, horizon scanning, and impact assessments. In the Hungarian case study, the impact of various interventions on soybean production was examined from very different economic and environmental perspectives.

The GLOBIOM model, developed by the International Institute for Applied Systems Analysis (IIASA), is a global economic model of land use that links agriculture, forestry, and bioenergy. It is used for land use and agricultural product projections, and a whole range of market outcomes (prices of agricultural and forestry products, international trade flows), as well as climate change impacts. It has also been specially developed for greenhouse gas emissions assessment. From the perspective of EU policies, it specialises in EU climate policies, which was the subject of the Czechia CS and proved to be relevant.

However, as agricultural policy issues have expanded into new areas in recent decades (such as climate change, biodiversity, and, most recently, social issues, animal welfare, and antibiotics), quantitative models cannot now cover all of the issues relevant to decision-makers given their currently available databases. There are different strategies to address new challenges. Analyses of social issues are often hindered by the lack of relevant household-level data, including all income sources. It follows that improved access to and coverage of farm households in micro-level datasets might address this deficiency. Also, for questions relating to antibiotics and pesticides, data availability hinders more complete coverage in modelling exercises. Animal welfare can also be addressed for specific questions with large-scale models (example: End of cages study for sows by CAPRI (Potori et al., 2023)), but full-scale coverage would require daunting data collection efforts. Sometimes a good pragmatic solution for a comprehensive assessment of the effects is to use several models in conjunction and link them together (as we will show in the integral case for The Netherlands).

The techniques of experimental economics focus on individual policy-relevant issues. It is important that the choice of method, data, and implementation is tailored to needs. We were unable to confirm this with our example, but in fact, the theory of experimental economics allows for a strong focus on individual issues, as demonstrated by numerous scientific publications in this field.

EXPERIENCE

Experience refers to the extent to which the tool has been used across MSs, in previous programming periods, or in other policies.

Evaluation Question: How experienced are MSs with this category of tools? Has the tool been used in previous programming periods of the CAP, or for other policies (e.g. climate, environmental etc.)?

In MSs where CSs were conducted, strategic planning was not supported by the selected type of quantitative tools, so these are only the first experiences of these MS. Impact assessment for the implementation of the CAP at MS level is not mandatory, so only a few MsS carry out similar demanding analyses at their own discretion. In the context of the CS in Czechia, the model had not been used by policy makers, but CzechGlobe has 5 years' experience with the tool funded by national projects for assessing climate change impacts and adaptation strategies for the agricultural sector.

It is therefore very important that impact assessment is also formally prescribed and that MSs are encouraged to use quantitative tools and thus to design agricultural policy in a more evidence-based manner.

Experimental economics, mainly due to reasons outlined above – high requirements in terms of financial, human and time resources – have heretofore rarely been practically applied in EU agricultural policymaking. These methods have only relatively recently begun to be applied to agricultural policy, to inform especially the design of agri-environmental measures. However, their influence on CSPs has typically been indirect, occurring through preparatory studies, pilots and evaluations (often through applied research conducted by national agricultural research institutes and universities) rather than explicit references in CSP documents. In contrast, experimental approaches are more firmly institutionalised in other EU policy domains, where dedicated behavioural units and clearer demand structures exist.

[5] https://web.jrc.ec.europa.eu/policy-model-inventory/explore/models/model-capri/policy-support/

Granularity

Granularity refers to the extent to which the tool can provide results at different level of spatial resolution (e.g. local, regional, national, and EU level).

Evaluation Question: Can the tools support analysis and policymaking at different levels, including local, regional, national and EU level? Do the tools provide insights at different geographical and time scales (e.g. results expressed in days, years, weeks)?

The dynamics of the CAPRI model are important in a time scale context. It is basically a comparative static partial equilibrium model, which means that it does not include dynamic relationships across time periods (except for LULUCF effects), while dynamics are more important in the GLOBIOM model. GLOBIOM is recursive-dynamic, meaning that each time point is solved separately and sequentially, and the results are then carried forward and constrain the next period decisions. The large-models do not aim to simulate future events, but to provide 'what if' scenarios following a holistic way of projecting the future of the land-based sectors. They work in the realm of possibilities, instead of the realm of probabilities of certain developments occurring or not occurring.

The GLOBIOM and CAPRI models cover the entire world. Smaller countries are grouped into market regions that may depend on the model version. CAPRI also has a global market model where some countries are represented individually or grouped in trade blocks. CAPRI is very detailed for the EU, covering NUTS 2 regions (around 280 regions). GLOBIOM's simulation units are even smaller than NUTS 2, but regional results are rarely analysed in detail. Comparative studies for Czechia, Hungary, and Spain have shown that both selected models can also be used effectively for questions relevant at the MS level. It is important to emphasise that al CSs used the full regional detail and thus provide a broader picture, even if the focus is solely on one MS. For the EU, a more detailed GLOBIOM version goes down to the country level, while a more aggregate version only identifies groups of countries ("EU_North").

Experimental economic methods have a clearly defined location to which the results apply, as they typically produce context-specific results. This reflects the behavioural nature of these analyses, which capture responses shaped by local specificities, including socio-economic and cultural conditions. As a result, findings are most directly applicable to the setting in which the experiment is conducted, while transfer to other contexts generally requires careful adaptation or replication. This was also the case in the French case study. Some techniques also allow for the prediction of future values over time.

Accessibility

The accessibility assessment aims to evaluate the extent to which the tools and their underlying methodologies, data, assumptions, and outputs are clear and accessible to stakeholders and users. The results of the evaluation are presented below.

Ease of use

Ease of use refers to the extent to which the tool is easy to learn and use.

Evaluation Question: Does the category of tools require prior training or experience, or does it require a great deal of time in order to be able to use it properly? Can these tools be handled by a wide range of professionals, or do they require specific professional skillsets in order to be implemented?

The maintenance and implementation of mathematical simulation models such as CAPRI and GLOBIOM is usually the responsibility of research institutes, as they have the necessary workforce and investment capacity to manage such tools. Both models used in the CSs are the result of many years of development at the highest scientific level by top modellers, and getting started with these models is a challenging endeavour. Special expertise is required to operate these models and translate policy changes into model parameter adjustments.

CSs have confirmed that the use of the proposed tools requires prior experience. Both models are generally quite complex, so researchers need to be well trained and experienced to be able to use and apply them. A prerequisite for CAPRI and GLOBIOM is that timely investment in research/modelling infrastructure is desirable to enable relatively easy application at the required time. Models of this type are hosted by one or more research institutes, but are used, maintained, and disseminated by a community of researchers. The CAPRI community is based in Bonn, Braunschweig, Madrid, Uppsala and JRC Seville and JRC Ispra, while GLOBIOM is based at IIASA in

Laxenburg (near Vienna). GLOBIOM is also expanding its network of developers, a good example being Czechglobe, which carried out a case study in Czechia.

Working in such research communities has the advantage of allowing researchers to share experiences and distribute the maintenance and investment costs associated with the models. Both case studies have shown that local research groups cannot operate the models or adapt them on their own; they need training and ongoing support from the core group of researchers behind both models. Local researchers, of course, benefit from the international research community, which offers them advice, assistance, and support when using the tool at the national level.

The CSs have shown that, with good external support, it is possible to develop national versions of both large models that can produce relevant studies and results in the context of specific issues in a MS. However, this is only possible if there is external support that takes on the bulk of the development work and coaches the local researchers so that they understand all of the tasks associated with the operation of the model and, if necessary, the steps required to carry out the modelling activities themselves.

For long-term use, a national research group should be developed that would work with the CAPRI or GLOBIOM model on an ongoing basis and would gradually become accustomed to model development and major model adjustments, which would also enable the performance of tasks for the needs of the relevant ministry. However, this is only the first condition for the expansion of a complex tool in a MS context. It is essential that such human capital investment has stable ongoing financing from the MS, because skills to use complex models can be lost or will not reach the minimum threshold required if staff members do not have the chance to accumulate knowledge over an extended period of time. In any case, this is a long-term process that goes beyond the scope of a two-year project facilitated by the project's CS approach. Therefore, the relatively successful implementation of CSs using CAPRI and GLOBIOM is only the first step towards implementing such models in a MS context.

The introduction of support for strategic planning using experimental economics methods also depends heavily on the development of this type of knowledge in the MSs and on the understanding of the client's needs and interests. As can be seen from the CS in France, which, while it cannot claim to have global relevance (it was carried out only with the regional government, and the communicator was not the largest national institute, INRAE), nevertheless shows how the modellers and their clients must be trained to cooperate and use such tools. Additional competencies and experience with similar research are certainly needed, as well as an understanding on the part of government officials of the use possibilities that exist with the models. This type of supply and demand cannot be established overnight or through a single project. Here, as in modelling, it is crucial that key government officials in responsible positions recognise the need for a more evidence-based approach to agricultural policy planning and implement it with the support of quantitative tools. This requires planned and long-term stable activities and investments on both sides with a clear goal of building this kind of support.

Tool Availability

Availability refers to the extent to which the tool is accessible, available to and usable by end users.

Evaluation Question: Are these tools openly available to everyone, and are their methodological and technical guidelines accessible to users? What are the main factors that limit their availability (e.g. costs, language, ease of use, available support, complexity, technical requirements)? Are there legal or patent-related limitations to their use?

Access to both models implemented in the CSs and their research communities is generally quite open. Nevertheless, report 1.3 (Tools4CAP, 2024) states that GLOBIOM is not available to end users in all MSs. Users interested in using the tool often have to sign a memorandum of understanding (MOU) to become a member of the modelling community and gain access to the tool, including the model code or data.

For CAPRI, the access to the stable release version of the model code and data is free to anybody, but successful application of the model will usually require significant support from the CAPRI core teams. There is then a certain reciprocity involved in order to progress the model's usage: the partner receiving the tool can benefit from the opportunity to use the tool and participate in the training provided. In return, the new national team is expected to contribute to the modelling consortium. This contribution can be financial, for example through projects, or in kind, by providing data on national conditions and contributing to model improvements. For CAPRI, a MOU has sometimes been used to clarify and organise research plans below the level of fixed contracts. But often such understandings simply evolve as a result of some collaboration, without any written agreement.

Complex mathematical models such as CAPRI and GLOBIOM have fixed costs and require constant extensions, updates, and improvements, which are expensive, especially since such models typically cover a wide range of countries and hence support cannot be granted to all interested individuals without considering some form of reciprocity.

Both CAPRI CSs showed that such access can be quickly provided if there is mutual interest and available project funds. After the end of Tools4CAP, however, things will come to a standstill unless memoranda of cooperation are signed and mutual interest is demonstrated. Given the results of the CSs, this would be a missed opportunity. In the case of Czechia, Czechglobe has been using GLOBIOM for over 5 years prior to this project, so there is opportunity that the work will expand to include direct studies and impact assessments for government needs.

Experimental economics methods are broadly shared analytical methods, part of open science, and therefore accessible to anyone able to implement them. The accessibility of the method is thus linked to the availability of appropriate research infrastructure, human resources, and the development of an evidence-based approach at the ministry level. The CS has clearly demonstrated this.

Efficiency

The assessment of efficiency aims to evaluate the extent to which the tools are able to achieve desired outcomes given the resources available, including time, human resources, and costs. The results of the evaluation are presented below.

Time Efficiency

Time efficiency refers to the extent to which the tool is able to achieve its outcomes within a reasonable amount of time. It includes the time necessary for its set up and for the implementation.

Evaluation Question: How much time does it take to set up this category of tools (i.e. time necessary to organise and prepare the implementation)? And how much time does it take to produce useful outputs from the tool (i.e. time necessary to implement the tool until the achievement of the intended outputs)?

CSs have shown that using large and complex simulation models such as CAPRI and GLOBIOM is time-consuming and requires a lot of time and effort. Establishing a new modelling team in an institution that has not previously used such models can take several years to set up and implement, even with significant assistance from the main developers. National researchers will also need sufficient capacity to be able to become fully independent. However, as the system is used by entire research networks, technical issues can be addressed to experts in the specific field. As already mentioned, this type of operation requires long-term cooperation between external model developers, the national research group involved, and the relevant ministry, which commissions this type of work and studies.

The CAPRI model has been operational since about 2000, while GLOBIOM was established a decade later, so many years of investments have been used in the establishment of these systems. With this foundation, the response time for some issues at a national level can be less than a year for the core team, but it typically takes at least three years for new entrants to become relatively autonomous model developers and address issues in relative independence of the core teams.

Experimental economics methods can vary in terms of time requirements. It all depends on the question posed and the method chosen. The French CS has shown that it is possible to carry out a simple assessment of the intentions and opinions of potential beneficiaries of the measure in a relatively short time. If the MS had the capacity and infrastructure, testing in experimental economics laboratories, where controlled experiments (combining economics, psychology, and sometimes computer simulations) are carried out to study human behaviour, the process could also be relatively quick. Larger DCE analyses, which require at least 300 or more survey responses, are much more time-consuming. The minimum time from the design of the study to the dissemination of results is at least one year. To begin with, it would be sufficient to continuously check the expectations and approaches of the potential beneficiaries of planned measures, using simpler methods such as focus groups or interviews. This practice exists in only a few MSs.

Cost-effectiveness

Cost-effectiveness refers to the extent to which the tool is able to achieve its outcomes with a reasonable budget. It includes the budget necessary for patent subscriptions, maintenance, and implementation.

Evaluation Question: Is implementing these tools expensive? How much budget can be necessary for the implementation of the tool? Is acquiring and/or maintaining these tools expensive? How much funding is necessary for the maintenance of the tool or for subscriptions for patent-based tools every year?

Large models such as GLOBIOM and CAPRI are essential for analysing a country within the EU and international context, particularly for studying market and trade interactions and climate-trade relations in scenarios such as trade reforms or geopolitical shocks. Although these models are very resource-intensive, their use for these specific research questions is considered cost-effective. Sometimes (but not always) a fee is required to use, while reciprocity is usually a condition for participation. Based on the results of the study (Tools4CAP, 2024), it is not possible to accurately estimate the required budget and the costs of use and maintenance, as these models are often designed for specific research projects that enable the development and use of models. However, from a cost perspective, we estimate that at least 1 FTE resource at the national level and 0.5 FTE in the core group are needed to carry out such impact assessments, for at least 1 year. A rough estimate would be around 100,000 euro per year (depending on the salary level in the MS). At this level (perhaps even higher), this would also be the cost of maintaining such a model at MS level. This, of course, means permanent project funding. Designing a specific national scenario that does not require complicated code adjustments by the core team of CAPRI or GLOBIOM may be cheaper, as this may typically be handled with an effort of less than 0.5 FTE. But then national decision-makers would remain dependent on the outsourcing solution. Furthermore, capacities of the core teams are often a bottleneck.

The costs of implementing the experimental economics method can vary greatly. They depend on the choice of data collection method and the complexity of the data processing and analysis (e.g. DCE). Simple statistical analyses can be performed for a few thousand euro, while more complex projects can cost several hundred thousand euro. In any case, these are research or expert projects carried out by a research group selected by the (public) client. Such a group does not arise out of nowhere; long-term investments in this type of research are necessary.

Human Resources Efficiency

Human resource efficiency refers to the extent to which the tool is able to achieve its outcomes involving a reasonable number of human resources.

Evaluation Question: How many people are necessary for the implementation of the tool and the analysis of results? How many different human resource profiles are needed for successful implementation of the tool?

For GLOBIOM and CAPRI, the number of people required less important than their capacity, knowledge, and skills to work with this tool. In CSs, the models used are usually supported by associated international research communities. These communities are responsible for developing and maintaining the models, organising conferences, collaborating on projects, and providing the necessary training. Effective use of the models requires the involvement of researchers in the relevant communities, as these applications typically require teamwork. Participation in such research communities offers the advantage of access to and utilisation of experience and expertise relevant to the project.

A one-off project-based action, as carried out in the CSs, does not require significant staff involvement (see above: 1.5 FTE). Otherwise, of course, regular updating of the CAPRI and GLOBIOM systems requires constant involvement of researchers and, given the reality of the situation in research institutions, this is limited due to a lack of manpower, resulting in less frequent major revisions than would be desirable. However, most database or module updates are handled by the core teams and thereby code developers and maintainers with extensive expertise. The role of national model users is to check the national data and policy specifications and to update and modify them for scenarios of national interest. This also requires at least a basic understanding of modelling systems and hence human expertise in a very specific area that needs to be accumulated over a number of years rather than months.

Regarding experimental economics, it is not possible to precisely define the need for personnel. Simple analyses (descriptive and simple inferential statistics, as was the case in the French CS) can be carried out by a small research team or expert group with experience in social science research. Any serious work in experimental economics requires experienced and well-trained staff. As a rule, these are groups of 1 to 5 researchers who will habitually also be working on other projects. It is important to maintain constant cooperation with clients in the agricultural field to understand their needs and be able to translate them into research questions, methods and measures.

2.1.3. Key findings

The report evaluates the use of quantitative tools, the CAPRI model, GLOBIOM model, and experimental economics methods, in supporting the strategic planning and assessing the impact of CAP options at the MS level. It focuses on four case studies: Czechia, Hungary, Spain, and France. The analysis is structured around key TEF criteria: *accuracy*, *reliability*, *applicability*, *accessibility*, and *efficiency*.

Accuracy: Models such as CAPRI and GLOBIOM capture complex agricultural systems, but require years of development, expertise, and continuous dialogue with policymakers. They are based on robust, harmonised datasets (e.g. EUROSTAT, FAOSTAT, FADN) and maintain internal consistency in the spatial and market dimensions. Despite their solid methodological foundations, there are limitations. Their application at MS level generally requires additional data verification and specific model adjustments depending on the topic. Although these models cover all EU MSs, there are often insufficient resources for in-depth tailoring of the models to best reflect national-level data and associated parameters. This detailed data verification, and sometimes additional calibration and sensitivity testing, is needed to ensure that the CSs are successful. Experimental economic methods rely on tailored data collection (surveys, interviews), but time constraints and methodological expertise often limit their applicability in policy-related applications.

Reliability: Full replicability of scenario analysis across various models is difficult due to their different theoretical bases and data sources. However, consistency in direction and theoretical consistency can generally be achieved, having been tested in several projects involving model comparisons. CAPRI and GLOBIOM provide public documentation, training materials, and open access to basic model versions, although advanced expertise on the part of the researcher is required to understand their structure. Experimental economic methods are transparent by nature, supported by scientific rigour and open science principles. The methods follow standards of replicability in theory, but practical constraints (e.g., small sample sizes) may reduce reliability.

Applicability: CAPRI and GLOBIOM support multi-objective scenario analyses, covering economic, environmental, and trade impacts. These models are widely used for the analysis of CAP reforms, climate policy, and trade scenarios. However, emerging policy areas (social issues, animal welfare, antibiotics) require the collection of extensive new datasets and extensions to these models. Sometimes the integration of multiple models with different databases and different approaches map help in achieving more comprehensive policy assessments. Experimental economics is excellent as a tool to address specific, targeted policy questions, but this requires tailored design and stakeholder involvement.

Accessibility: Complex models require specialised training, long-term investment, and integration into international research communities (CAPRI in Bonn/JRC; GLOBIOM at IIASA). Experimental economics methods are conceptually accessible, but they depend on the availability of national research infrastructure and access to government support.

Effectiveness: Setting up CAPRI/GLOBIOM at the national level can take several years and requires sustained funding (~€100,000/year) and expert teams (at least 1.5 FTE). Experimental economics varies in its resource requirements: simple analyses cost at least 10.000 euro, while complex projects can reach several hundred thousand euro. Both approaches require highly skilled staff and long-term cooperation between research institutions and ministries.

We can conclude that national adaptation of EU-level models, such as the two tested in the CSs, requires external support, training, and stable funding at the national level. There are many conditions that must be met for such a complex model to be used effectively at the MS level. The willingness and ability of domestic researchers to decide to invest in such a topic is crucial, but there is a long way to go before they can use such tools themselves. Lower entry barriers exist when using national results derived from existing larger scale analyses, but often national decision-makers would prefer tailored scenarios specifically relating to their own MS. Therefore, it is unlikely that such tools will be used more widely at the MS level in the future.

If there is interest in the research community and at ministerial level, we recommend that potential practitioners and end-users consider opting for less complex models, such as AGMEMOD, if those models also answer their questions, or that they envisage the construction of simpler software models targeting the level of agricultural holdings, either based on the FADN data principle or on a selection of typical farms. The results of the models are also easy to communicate to the public and the models can more easily accommodate new detailed requirements for beneficiaries of CAP funds. For a comprehensive policy assessment, the integrated use of multiple simpler models or engagement with modelling teams responsible for large-scale models may be recommended. The principle is clear: simple or narrowly defined questions are best addressed with simpler tools, but if questions become

complex and relevant for several topic areas, more complex model combinations or large-scale systems are pertinent. In any case, formalising impact assessment requirements at the MS level would encourage evidence-based policymaking using models.

Experimental economics offers valuable insights into farmer behaviour, but it needs institutional capacity building and stakeholder involvement. The CS conducted revealed an interesting phenomenon: despite the scientific community emphasising the usefulness and importance of this type of methodology for more evidence-based policymaking, the clients (end users of the results) of these techniques do not make use of the results of these models frequently or commission scenario analysis using these models on a regular basis. The usefulness of such approaches has already been demonstrated, and their results are being transferred into the debate on measures (at least in research circles), but their direct application is very limited (if it exists at all, it is more incidental and through targeted research projects). However, the experimental economics CS clearly showed that even very simple surveys can improve the understanding of the intentions and capacity of beneficiaries to engage with measures and help decision-makers to better design measures.

2.2. Qualitative and participatory tools

2.2.1. Summary of the Case study reports

Case study Germany – Using the IOI matrix to assess coherence between eco-schemes and AECM in Hesse

Germany's federal structure creates complexity in CAP governance: national authorities design eco-schemes, while regional states implement agri-environment-climate measures (AECM). Ensuring coherence between these instruments is crucial for achieving environmental and climate goals and avoiding duplication or conflicting incentives. The case study in Hesse tested the IOI (Intervention–Outcomes–Impact) matrix as a tool to analyse how measures under Pillar I and Pillar II collectively contribute to policy objectives. The aim was to evaluate coherence, identify gaps or overlaps, and explore how the tool could support multi-level coordination.

The IOI matrix maps interventions against expected outputs and impacts, highlighting their contribution to specific CAP objectives such as water protection, biodiversity conservation and climate mitigation. Through desk research and expert consultations, researchers populated the matrix for selected eco-schemes and AECM in Hesse. The exercise facilitated structured dialogue with the regional ministry and revealed interactions of national and regional funding streams in pursuing environmental and climate goals.

The results show that the IOI matrix is a useful framework for visualising coherence across instruments and levels of governance in a concrete way. It increases transparency, supports informed discussion, and provides a basis for adjusting programme design. The method also revealed several weaknesses: data gaps, limited availability of quantified impact estimates, time constraints, and the reliance on expert judgement. Engagement from administrative staff was constrained by both time pressures and frequent turnover, reducing opportunities for in-depth validation. Furthermore, broad stakeholder participation was not feasible due to limited time and administrative capacity, forcing the CS team to rely on expert opinion. Nonetheless, the matrix proved valuable as a communication tool that bridges scientific analysis and administrative decision-making.

The study recommends integrating the IOI matrix into regular planning and review processes for both eco-schemes and AECM. Future work should focus on strengthening data foundations, improving impact quantification, and building administrative capacity to apply the tool independently. Wider stakeholder involvement could enhance legitimacy and lead to more holistic discussions about policy coherence. If systematically embedded, the IOI matrix can support evidence-based adjustments to CAP programmes and improve coordination in Germany's multi-level governance context.

Case study Ireland – Developing Intervention Logic (IL) to improve CAP monitoring and evaluation

The Irish CS was characterised by a need for better linkages between certain interventions, outputs and outcomes such as farmers' livelihoods, farm landscapes and food security. Existing PMEF indicators were considered insufficient for national evaluation needs, and the Department of Agriculture, Food and the Marine (DAFM) sought to strengthen the evidence base for monitoring and evaluation, as well as a basis for discussion with stakeholders.

The aim of the case study was to develop a detailed Intervention Logic (IL) that maps causal pathways from CAP measures to expected outcomes, while identifying additional indicators relevant to Irish priorities.

The study combined a review of the literature, expert workshops, a Delphi survey, and analysis of National Farm Survey (FADN/FSDN) sources to construct an IL covering direct payments, eco-schemes and payments for areas with natural constraints. In addition to a description of the expected outcomes of the assessed interventions, the process helped to identify and refine feasible indicators and potential issues related to their estimation, as well as data gaps to be addressed in order to make the indicators robust.

Key findings show that IL can significantly enhance the coherence and transparency of CAP planning, implementation and evaluation, as well as inform subsequent policy cycles. It can also be used as a communication tool and, by identifying the expected outcomes of each DP intervention, facilitates the selection of key performance indicators to assess the effectiveness of interventions. At the time of reporting, the elaborated IL was still to be validated by DAFM personnel.

In terms of recommendations, an effective IL should be developed through an evidence-based, participatory process. Work should begin with a review of scientific and policy literature, followed by workshops with policy officials and thematic experts to identify expected outcomes and impacts. A draft IL should then be reviewed by relevant stakeholders and finalised for public use. The IL should be treated as a living document, regularly updated and used to support monitoring, evaluation, indicator selection and future CSP design. At least one person with experience in policy planning and evaluation should guide the process. The Irish CS indicates the value of extending IL development to all CSP interventions and the benefit of strengthening environmental and social indicators to support a more comprehensive assessment across CAP objectives.

Case study Poland – Participatory prioritisation of needs using the Constrained Cumulative Voting (CV) method

Poland's experience during the preparation of the CAP Strategic Plan 2023–2027 revealed challenges in prioritising needs in a transparent, inclusive and evidence-based manner. The CSP development process was characterised by a lack of time for open and transparent discussion, resulting in a bias towards priorities reflective of historical preferences. To address this deficiency, the Polish CS tested a participatory prioritisation method based on Constrained Cumulative Voting (CV), a structured point-allocation approach that is perhaps best described as a quasi-quantitative method. The objective was to determine whether structured participation could improve the understanding of trade-offs in CAP decision-making and open up the policy debate regarding the use of tools in CSP design and implementation.

The CV method was applied during a facilitated workshop with representatives based on the "quadruple helix": administration, business organisations, academia and civil society. Stakeholders were tasked with allocating 100 points among a list of identified needs, with lower and upper bounds to prevent dominance by any single preference. This approach forced participants to recognise trade-offs and required them to give attention to multiple priorities. The exercise proved relatively easy to understand and implement. However, actually prioritising needs proved difficult because results were highly sensitive to available criteria that could be selected, the point-allocation rules, stakeholder composition, and the way the (large number of) needs were formulated. Complex or multi-dimensional needs made it hard for participants to assess relevance consistently, indicating the importance of clear, well-structured definitions of needs.

The findings show that a structured participatory method can enhance transparency and stakeholder ownership, while reducing the perception that prioritisation is politically predetermined by the ministry's preferences. At the same time, the workshop revealed tensions between different priorities, which the method helped surface and document. Limitations with regard to the representativeness of the findings include the relatively small and (unavoidably) non-representative group of participants, time constraints, and a lack of political interest in the ministry. More generally, a key risk when using participatory tools is that they can be perceived to have been used symbolically (or performatively) rather than being meaningfully linked to actual decision processes.

The CS authors recommend that prioritisation be kept simple, transparent and inclusive, as overly complex tools risk bias perceptions and reduce stakeholder trust. Needs must be clearly formulated, avoiding multi-dimensional descriptions that hinder consistent assessment. The criteria used for prioritisation should be carefully selected and tested with representatives of all stakeholder groups to ensure they are well understood and relevant. Limiting the number of points that can be allocated to a single need can help achieve a more balanced allocation of scores. Preparation and facilitation should be carried out by at least two people, with broader stakeholder testing before

implementation. Although the tool is inexpensive, time requirements increase with the number of actors involved. Future processes should consider hybrid prioritisation methods (e.g., Analytical Hierarchy Process (AHP) or the Entropy Weight Method (EWM) for handling complex objectives. Finally, further research is recommended to tailor prioritisation exercises to different CSP stakeholder groups and validate the proposed improvements.

2.2.2. TEF analysis

ACCURACY

Accuracy analysis considers how successful is the tool in offering results that reflect reality without any distortion or omission in the results. The results of the evaluation are presented below.

DATA ACCURACY AND VALIDITY

Data accuracy and validity refers to the extent to which data input and its sources are traceable and validated.

Evaluation Question: To what extent do the tools rely on well-founded evidence (e.g. the sources of data used during the implementation of the tool are known to be reliable)? Is it possible to easily trace the information used for supporting the policy choices back to its original source? Are measures in place to avoid mistakes and redundancy and to ensure consistency in the data or information used to support the policy choices (e.g. to verify the quality and reliability of the information)? Are results triangulated, corroborated, or substantiated by other data sources?

For qualitative tools, data is less about large numerical datasets and more about documented evidence (including written sources), expert knowledge and stakeholder judgments. The IOI matrix in Germany relied primarily on literature review, programme documents and expert assessments of how instruments contribute to objectives. When populated carefully, the sources are highly traceable: each cell of the matrix can be linked back to a specific measure, regulation or official document, and the reasoning for the coherence score can be documented explicitly. This is broadly in line with other coherence-matrix approaches used in environmental and Sustainable Development Goal (SDG) policy analysis, which rely on structured expert elicitation and transparent documentation of assumptions (Mortelmans et al., 2020).

In Ireland, the Intervention Logic (IL) builds directly on CAP Regulations, the CSP, PMEF indicators and FADN/national statistics, as well as additional desk research of scientific sources. Logic models are widely recommended to clarify causal chains and make underlying assumptions explicit; they are considered good practice in evaluation and CAP guidance precisely because they force evaluators to link data and indicators to specific steps in the causal pathway (MacLeod, 2016).

In Poland, the Constrained Cumulative Voting (CV) method produces prioritisation scores that are immediately traceable to individual stakeholder allocations, though the method can be conducted anonymously. The method is well documented in decision-making and prioritisation literature and is considered a relatively simple, ratio-scale way of capturing preferences (Brams & Fishburn, 2002). Accuracy here depends mainly on clarity of the list of needs, the quality of briefing materials, and the careful recording of votes. The method is highly sensitive to the selection of participants, as CV reflects only the preferences of the individuals involved. If the goal is to represent the preferences of the broader population rather than specific stakeholder groups, a non-representative selection of participants will inevitably limit the generalisability of the results. The tool itself does not introduce bias, but the composition of the participants can influence outcomes in ways that should be acknowledged when interpreting results.

Across all three cases and participatory tools in general, validity is strengthened when:

- the selection of documents and stakeholders is systematic and transparent,
- assumptions behind scores (IOI, IL causality, CV thresholds) are made explicit, and
- results are cross-checked with other evidence (e.g. monitoring data, previous evaluations).

However, none of the tools inherently guarantees accuracy; it depends strongly on how rigorously the input information (from literature sources, expert opinion or stakeholder preference elicitation) is prepared, documented and triangulated.

Tool Representativeness

The capacity of a tool to accurately and comprehensively capture or depict a system, dataset, or phenomenon. It combines the notions of accuracy and coverage in its depiction.

Evaluation Question: To what extent do these tools present a high degree of sensitivity (i.e. how much are the results of the tools influenced by changes in the input data)? Are the tools able to capture the diverse facets, variations, or elements within the system or dataset it seeks to represent?

Qualitative methodologies rarely strive for full population representativeness, as this is not their primary aim (Maxwell, 2012). Rather, the aim is conceptual representativeness, often relying on purposive sampling, which strives for relevance and seeks to reflect a diversity of viewpoints and appropriate coverage of the phenomenon in question. For qualitative tools, the concept of representativeness refers to whether the tools faithfully capture the diversity of objectives, measures and perspectives relevant to CAP strategic planning or a selected facet of it. The German IOI matrix can, in principle, cover the full range of eco-schemes and AECM and all specific objectives, offering high "conceptual coverage" of the policy mix. But the sensitivity and degree of representativeness of the matrix depend on the capacity of the expert group and the quality of available evidence and policy documentation. Similar policy-coherence matrices in other fields are recognised as useful for mapping interactions, but are known to be sensitive to who is selected to complete them and how criteria are defined (Mortelmans et al., 2020).

In Poland, the CV method can represent stakeholder diversity if the "quadruple helix" (administration, business, science, civil society) is adequately represented. This means that exceptional care must be taken in selecting the appropriate stakeholders, as this is crucial for achieving a high level of representativeness. As for sensitivity, a constrained design (upper and lower bounds) can be used to prevent any single interest from dominating, a feature also emphasised in the cumulative-voting literature (Cagliero et al., 2022). Furthermore, the method is also effective in capturing changes in stakeholder preferences as a reflection either of stakeholder selection or actual changes in preferences due to contextual changes. The method is also very sensitive to minor methodological tweaks, such as e.g. whether 'voters' can see each other's votes, whether they are allowed to discuss amongst each other, etc. But representativeness is still limited when the number of participants is small, topically or spatially (e.g.) regionally unbalanced.

The Irish IL case is representative mainly in content: it attempts to include all relevant interventions and outcomes along the results chain, not all possible actors. It is well aligned with international practice where logic models are used to structure complex policies with multiple objectives (MacLeod, 2016). IL can be considered to somewhat abstract in nature and is not easily amenable to stakeholder involvement.

Overall, these, and other qualitative tools, are good at representing complexity and interlinkages, but their representativeness hinges on careful coverage of measures and inclusive stakeholder selection rather than statistical sampling.

Reliability

The reliability assessment aims at measuring the consistency and stability of the results obtained by the application of the selected tools. This criterion measures the extent to which the process of implementation of the tools produces consistent findings when repeated. The results of the evaluation are presented below.

Replicability

Replicability refers to the extent to which the tool's results are (or can be expected to be) consistent across studies aimed at answering the same question, either by replicating the same approach or using new data or methods.

Evaluation Question: Can the results of these tools be replicated successfully by others or through different approaches? Do the results remain consistent with other studies and approaches?

The procedures for IOI matrices, IL development and CV workshops are highly replicable. However, the actual numerical results (scores, rankings) are highly context-dependent and are not expected to be identical if actors or contexts change. In Germany, another team could replicate the IOI procedure (same template, same coding rules) but might arrive at slightly different coherence scores if the expert composition, available evidence or interpretation differs; this is typical of policy-coherence matrices (Mortelmans et al., 2020). One approach to mitigate this risk, if the need for replicability is high, is to ensure enough time for the process, include a relatively high number of stakeholders, and complement the method with iterative consensus-building tools, such as the Delphi method (Linstone & Turoff, 1975).

As the Polish CS reflects, CV itself is a rather standardised method: if the same group repeated the exercise with the same information and constraints, one would expect broadly similar prioritisation patterns, which is consistent with studies on cumulative voting in other domains (e.g. Marta-Pedroso et al., 2018). However, because preferences can change with new information or political context, divergence over time does not signal a lack of reliability, but is instead reflective of actual changes in preferences.

The IL, as applied in Ireland, is entirely replicable as a conceptual framework: different teams could reconstruct the logic, but choices about the level of detail, grouping of interventions and selection of indicators and the evidence base to substantiate different logical links will inevitably include a subjective element. Logic-model literature stresses that these tools are meant to be living documents revised as understanding evolves, not fixed "truths" to be reproduced identically (MacLeod, 2016).

Thus, reliability for qualitative tools should be judged by clarity and repeatability of the method and consistency of high-level patterns across applications, rather than exact numerical reproducibility of results.

Transparency

Transparency refers to the extent to which information concerning underlying methodologies, data, and results of the tool are available and accessible, as well as the extent to which the tool is open to external review.

Evaluation Question: is the information concerning underlying methodologies, data, and results of the tool available and accessible, and is the tool accessible for external review?

All three qualitative tools score relatively high on transparency when applied according to good practice. IOI matrices can be published as simple tables showing which measures contribute to which objectives, with scoring rules and sources described. Similar policy-matrix methods used in environmental and SDG coherence work are explicitly designed to make assumptions and trade-offs explicit and visible to decision-makers and external reviewers (Mortelmans et al., 2020).

Cumulative voting is also very straightforward to explain: stakeholders see the list of needs, the voting rules and the aggregated results. Depending on the modality of application (which, in turn, depends on the purpose), they may even be able to see each other's votes, stimulating discussion, as also reflected by a part of the Slovenian CS aimed at prioritising needs. The method has been recommended in participatory decision-making because it produces easily communicable outputs (point allocations and rankings) and invites discussion of underlying reasons (Cagliero et al., 2022).

Intervention logic, by definition, is a (visual) device to expose how policymakers believe change will happen; it is widely promoted in CAP evaluation guidance and broader evaluation literature precisely because it supports transparent planning, indicator selection and communication (MacLeod, 2016). In theory, the better the intervention logic, the better the result and legitimacy of the policy measure(s).

The main transparency risk lies not in the tools themselves but in the quality of related documentation: if the evidence base, workshop minutes, scoring results or IL diagrams are not adequately written up and shared in full, including a record of contentious points and the methods applied to resolve them (if any), the process becomes opaque. Where documentation is good, these qualitative tools are highly accessible and reviewable.

Applicability

The applicability assessment aims at measuring the extent to which the tool can be used to different contexts and pursue different objectives. The results of the evaluation are presented below.

Multi-functionality

Multi-functionality refers to the extent to which the tool pursues several objectives and policy-making steps at the same time, (i.e., the number of policy objectives or policy-making steps that can be addressed by the tool), as well as the extent to which the underlying assumptions of the tool can be easily changed to address other policymaking steps or other objectives.

Evaluation Question: To what extent can these tools be used and/or adjusted to address different policy objectives and steps of the policymaking process?

The three qualitative tools applied in the case studies exhibit strong multi-functionality. IOI matrices can be used for ex-ante design (checking whether a planned mix of measures adequately addresses the full set of objectives), mid-

term review (identifying gaps and overlaps between policy measures) and ex-post evaluation (interpreting results in the context of an entire intervention mix or assessing the quality of the initial intervention logic by comparing expected and actual results of policy interventions). IOI matrices are adaptable to different policy areas and levels of aggregation, as shown by their use in environmental policy coherence and SDG analysis (Mortelmans et al., 2020).

The Irish IL explicitly serves multiple design and monitoring steps: needs assessment, design of interventions, indicator selection, monitoring/evaluation and communication of the CAP Strategic Plan. In fact, logic models are recognised in evaluation theory as highly versatile "backbone tools" for planning, monitoring and learning (MacLeod, 2016).

The Polish CV tool is primarily focused on prioritisation, which is relevant at several policymaking stages: selecting key needs, focusing evaluation questions, or designing targeted measures. Cumulative voting has been used in many decision-making contexts (from requirements engineering to participatory budgeting), which shows that it can be adapted to different policy domains and objectives. It can also be applied at different levels of abstraction, which means that it can be used for a range of purposes, from weighing high-level objectives to prioritising measures or even allocating funding.

Overall, the qualitative tools are flexible and can be used at different stages of the policy cycle or extended to other sectors or even, such as in the case of the IOI matrix, applied to improve policy coherence between different policy domains.

Experience

Experience refers to the extent to which the tool has been used across MSs, in previous programming periods, or in other policies.

Evaluation Question: How experienced are MSs with this group of tools? Have they been used in previous CAP programming periods or for other policies (e.g. climate, environmental etc.)? How many MSs have used the tool in this CAP programming period?

Internationally, logic models/intervention logic are well established in public policy and evaluation practice (Baehler, 2007; Matthews, 2015) and explicitly embedded in the CAP's monitoring and evaluation framework (Cagliero et al., 2021; MacLeod, 2016). Cumulative voting has a long history in corporate governance and decision theory and is recognised in the academic literature as a structured method for capturing preference intensity in participatory or multi-stakeholder settings (Brams & Fishburn, 2002). Policy-coherence matrices similar to the IOI matrix have been used in environmental agreements, SDG interactions and sector–environment policy studies (Mortelmans et al., 2020).

MSs generally have long-standing experience with qualitative and stakeholder-engagement tools, stemming from EU cohesion, environmental, and rural development policies where participatory approaches have been required for decades. Stakeholder engagement became obligatory for CAP Rural Development in 2014–2020 and for the entire CAP Strategic Plan in 2023–2027, meaning all MSs used qualitative tools in preparing their current CSPs. However, while workshops, consultations, and expert groups are widely applied, the use of more structured qualitative methods – such as prioritisation tools, coherence matrices, or systematic intervention-logic development – varies considerably across MSs, with only a subset employing them in a methodical and repeatable way. Formally, all MSs have been required to elaborate an intervention logic, although this has been applied to different degrees of quality. This highlights an important issue of differences in administrative capacity between MSs, which however does not necessarily correlate with the length of EU membership and experience with EU policies (European Commission, 2021). In addition, formal application of stakeholder engagement procedures has not always necessarily been reflected in final policy.

The specific combinations used in these case studies (IOI for eco-schemes vs AECM in Hesse, CV for CAP needs in Poland, IL for direct payments in Ireland) are relatively new and rather experimental. The CSs therefore represent early, but promising, applications of well-known methods to CAP Strategic Plan design and evaluation. The IOI matrix, in particular, has rarely been applied in MS, especially for ex-ante evaluations, showing a high level of potential for improving policy coherence.

Granularity

Granularity refers to the extent to which the tool can provide results at different level of spatial resolution (e.g. local, regional, national, and EU level).

Evaluation Question: Can the tools support analysis and policymaking at different levels, including local, regional, national and EU level? Do they provide insights at different geographical and temporal scales (e.g. results expressed in days, years, weeks)?

The qualitative tools offer substantial flexibility in spatial granularity, though they tend to be most effective at strategic scales. The IOI matrix can be applied at EU, national or regional levels and can zoom in on specific objectives while retaining a view of the overall policy mix. The Intervention Logic (IL) can be developed for an entire CSP, for specific intervention families, or for regional or thematic sub-systems. The (Constrained) Cumulative Voting (CV) method is typically applied at national or regional level, but can also be implemented in local or thematic settings.

In terms of temporal granularity, qualitative tools operate at relatively coarse time scales. They are suited to multi-year planning, major revisions, or periodic strategic reviews rather than short-term or high-frequency dynamics. They generally underpin "once per programming period" exercises rather than continuous or annual monitoring.

Additionally, these tools also offer flexibility in topical granularity: they can address broad policy architectures (e.g., the entire CSP), specific thematic domains (e.g., biodiversity, AKIS – Agricultural knowledge and innovation system), or narrowly defined intervention areas, depending on evaluators' needs and the scope of the assessment.

Accessibility

The accessibility assessment aims to evaluate the extent to which the tools and their underlying methodologies, data, assumptions, and outputs are clear and accessible to stakeholders and users. The results of the evaluation are presented below.

Ease of use

Ease of use refers to the extent to which the tool is easy to learn and use.

Evaluation Question: Does the group of tools require a great deal of prior training, experience or time to apply properly? Can they be handled by a wide range of professionals, or do they require specific professional profiles to be implemented?

Compared to complex simulation models, these qualitative and stakeholder engagement tools are relatively easy to understand for non-technical stakeholders:

- IOI matrices are simple tables with scores and comments.
- IL diagrams use intuitive boxes and arrows.
- CV uses a straightforward "X-points" allocation rule.

In the CSs, the main constraints that were reported were, therefore, not issues of technical complexity, but lack of time to organise workshops, availability of staff and stakeholders, and competing priorities in administrations.

That said, high-quality application of these qualitative tools still requires facilitation skills, basic evaluation literacy and time to prepare materials, which may also be copious and/or complex, depending on the topic. Guidance on logic models and intervention logic is now embedded in EU and national evaluation support materials, which flattens the learning curve (MacLeod, 2016). Cumulative voting is widely recognised in decision-making literature as easy to learn and implement, including when different variations are applied. However, as demonstrated in the Polish CS, effective application – especially for policy communication and intervention design – requires careful attention to how issues and options are formulated. If needs or concepts are unclear, overly complex, or insufficiently distinct, participants may struggle to allocate points meaningfully, leading to random or protest voting. This reduces the usefulness of the results and can undermine the perceived legitimacy of the process.

Tool Availability

Availability refers to the extent to which the tool is accessible, available to and usable by end users.

Evaluation Question: Are these tools openly available to everyone, and their methodological and technical guidelines accessible to users? What are the main factors that limit its availability (e.g. costs, language, ease of use, available support, complexity, technical requirements)? Are there legal or patent-related limitations to the use?

Unlike many models, these qualitative tools are not proprietary software; they are general methods described in various handbooks, academic literature and CAP guidance. Thus, legal and cost constraints are minimal; the main limiting factors are organisational capacity (people, time) and the availability of facilitation expertise, rather than access to the tool itself. In practice, external guidance or experienced facilitation often plays a decisive role in whether these tools can be applied effectively, particularly where administrations have limited experience with participatory or evaluative methods.

In addition, the effective use of these tools depends heavily on sustained interest and participation from policymakers and stakeholders. Engagement literature and experience of all the three analysed CSs show that participation can decline (or even be absent) due to meeting fatigue, competing administrative priorities, unclear incentives for involvement, or perceptions that consultation processes lack influence (OECD, 2020). Where participation wanes – whether because of resource pressures, stakeholder overload or disillusionment, reduced institutional commitment, or the reliance on voluntary and unpaid participation – the tools remain formally "available" but cannot be meaningfully applied. This is particularly relevant where farmer or environmental representatives participate on a voluntary basis, which can limit their capacity to engage consistently, especially in smaller Member States where resources to support professional representation are more limited. Thus, practical availability is shaped not only by methodological openness, but also by the willingness and capacity of stakeholders to engage, which constitutes a critical and well-documented limiting factor.

Efficiency

The efficiency assessment aims to evaluate the extent to which the tools are able to achieve desired outcomes given the resources available, including time, human resources, and costs. The results of the evaluation are presented below.

Time Efficiency

Time efficiency refers to the extent to which the tool is able to achieve its outcomes within a reasonable amount of time. It includes the time necessary for its set up and for the implementation.

Evaluation Question: How much time does it take to set up this group of tools (i.e. time necessary to organise and prepare the implementation)? And how much time does it take to implement the tool until the achievement of the intended outputs?

Once the basic templates are prepared, the implementation phase of these tools is relatively quick:

- An IOI workshop can populate a matrix in 1 to 2 days, although preparing the list of measures and objectives requires desk-based work completed in advance. Developing a sufficiently strong evidence base to substantiate the hypothesised links between interventions, objectives and impacts may be challenging and time-consuming, depending on the availability of existing analyses. In more complex cases, compiling and reviewing relevant scientific and policy literature can take several weeks.
- An IL for a specific intervention area may take several workshops and some analytical follow-up, but it is still faster than building or calibrating a quantitative model. Developing a sufficiently strong evidence base to substantiate the hypothesised links between interventions, objectives and impacts may be challenging and time-consuming, depending on the availability of existing analyses. In more complex cases, compiling and reviewing relevant scientific and policy literature can take several weeks.
- A CV workshop can be completed in a few hours once participants and materials are in place, which in turn can be prepared in approximately a day. In theory, to ensure that the voting results are both well-informed and based on stakeholder preferences, implementers may opt for building a stronger evidence base with materials (key issues to be considered) either sent to stakeholders beforehand or presented to them before the actual voting procedure.

Additional bottlenecks observed in the CSs were coordination and availability: securing participation from ministries and stakeholders, aligning agendas, and synthesising results into documents that can be used in the CSP cycle – all within tight CAP deadlines.

Cost-effectiveness

Cost-effectiveness refers to the extent to which the tool is able to achieve its outcomes within a reasonable budget. It includes the budget necessary for patent subscriptions, maintenance, and implementation.

Evaluation Question: Is implementing these tools expensive? How much budget is required for the implementation of the tool? Is acquiring and/or maintaining these tools expensive? How much budget is required for the maintenance of the tool or subscription for patent-based tools every year?

Qualitative tools are low-cost in terms of licences and software and moderate in terms of staff time and facilitation. The main expenses are:

- staff time for preparation and synthesis;
- venue/online platform and logistics for workshops;
- possibly external facilitation or evaluation support.

Compared with complex modelling exercises, these costs are modest and more easily absorbed within existing evaluation or technical-assistance budgets. Given that IOI, IL and CV can all inform multiple steps of the policy cycle (design, review, communication), their cost–benefit ratio is favourable, especially for Member States with limited modelling capacity.

Human Resources Efficiency

Human resource efficiency refers to the extent to which the tool is able to achieve its outcomes involving a reasonable number of human resources.

Evaluation Question: How many people are necessary for the implementation of the tool and the analysis of results? How many different professionality are necessary for the good implementation of the tool?

Each tool can be run by a small core team (1–3 people) plus a group of participants:

- IOI: 1 to 2 analysts + a group of experts/administrators;
- IL: 1 to 3 evaluation/analysis staff + policy officers;
- CV: 1 to 2 facilitators + stakeholders.

These methods require more generalist policy and evaluation skills rather than (niche) modelling expertise, making them more accessible to ministries and agencies without strong modelling units. However, effectiveness depends on having at least one person with strong facilitation and evaluation experience who can design the process, guide discussion and ensure documentation quality.

Overall, the qualitative tools used in Germany, Poland and Ireland offer a relatively efficient way to introduce structure, evidence and participation into CAP strategic planning, especially where quantitative capacity is limited or where rapid, consultation-rich processes are needed to complement more data-intensive analyses.

2.2.3. Key findings

In this section, we summarise and highlight the key aspects emerging from the evaluation of qualitative tools applied in the case studies from Germany, Poland and Ireland, offering insight into qualitative and stakeholder engagement tools more broadly. These tools support evidence-based CAP decision-making by structuring reasoning, eliciting stakeholder preferences and clarifying intervention logic. They help policymakers and stakeholders understand and communicate trade-offs, coherence and causal pathways when designing or adjusting CAP Strategic Plans.

Based on the evaluation, the main results can be summarised according to each criterion as follows:

Accuracy: qualitative tools depend on expert judgement, stakeholder inputs and document analysis. When assumptions, scoring criteria and decision rules are clearly formulated and documented, they offer reasonably accurate representations of policy logic and stakeholder preferences. However, because they do not rely on statistical inference, accuracy can be affected by incomplete participation, uneven expertise or insufficient preparation of materials. In many MSs, certain policy areas may also have a weak evidence base, or stakeholders may differ significantly in their capacity, power or ability to articulate preferences. In such contexts, results may reflect the balance of influence within the participating group or prevailing administrative perspectives rather than the preferences of the broader polity, which limits the generalisability of the findings.

Reliability: the procedures (IOI scoring, voting rules, IL construction) are replicable, and repeating them produces consistent types of insights. However, the results themselves naturally vary with different participants or contexts.

Reliability is therefore linked to procedural stability rather than uniformity of output. Qualitative tools are well-suited for reflecting complexity and capturing directionality and patterns, but are not devised to reproduce identical quantitative results.

Applicability: qualitative tools are highly flexible and can support multiple steps of CAP policymaking: needs assessment, prioritisation, coherence analysis, intervention design and evaluation planning. Their conceptual simplicity allows adaptation to different MSs and policy areas. Their applicability is limited mainly by institutional engagement, quality of preparatory work and in some cases existing evidence, not by methodological constraints.

Accessibility: these tools require no specialised software or mathematical modelling capacity. They rely on facilitation, evaluation literacy and stakeholder engagement – skills that are broadly available in administrations. Their widespread use in public-policy evaluation makes them readily accessible. Limitations stem from time availability and other constraints related to the engagement of stakeholders and staff, rather than technical barriers or severe financial constraints.

Efficiency: qualitative tools can be quickly set up and conducted. Workshops can be executed within days, and IL diagrams can be developed iteratively with modest resources. Costs are low, mainly involving staff time and facilitation. This makes qualitative tools particularly suitable for CAP contexts with relatively tight budgets, deadlines and limited modelling capacity.

Overall, qualitative tools offer high transparency, flexibility and efficiency, enabling structured discussion, reflection and stakeholder involvement in CAP SP design. Their main limitations (which are at the same time their main strengths) relate to subjectivity, dependence on participant diversity and quality of documentation. They are strong complements to quantitative modelling and provide crucial insights for evidence-based policymaking. However, it cannot be overstated that the quality of their results strongly depends on the quality of the underlying process used to produce the results.

2.3. Monitoring tools

2.3.1. Summary of the Case study reports

Case study Greece – Testing the Policy Monitoring Tool (PMT) using Earth Observation and digital agriculture data

Greece faces persistent challenges in developing an integrated monitoring system for CAP implementation, due to gaps in data integration and a lack of tools available to policymakers. As the CAP 2023–2027 places strong emphasis on performance-based governance, the Greek CS sought tools capable of linking administrative data with Earth Observation (EO) and digital agriculture sources. The CS tested the Policy Monitoring Tool (PMT) as a prototype for providing accurate, timely, and actionable data to support regional policy goals and inform national policy adjustments. It was anticipated that the CS would demonstrate the tool's capacity to strengthen monitoring processes and familiarise authorities' personnel with data-driven methodologies.

The PMT, a tool grounded in the integration of different data sources, was tested using a combination of previously collected data and open datasets within mock-up scenarios to explore the feasibility of generating reliable and relevant indicators. Following an institutional crisis at the originally intended partner (OPEKEPE, the Greek agricultural paying agency), the study collaborated with ELGA, the Hellenic Agricultural Insurance Organization, to evaluate data quality and tool usability. Preliminary results demonstrated that the PMT can effectively combine disparate data sources, automate indicator generation and visualise trends relevant to CAP performance monitoring. It also highlighted substantial potential to reduce administrative burdens via replicability across regions based on data and governance frameworks in place.

However, limitations were significant. In addition to issues related to data availability and interoperability, a lack of institutional access and cooperation severely hindered the CS's progression. To finalise the CS, its coordinators were forced to resort to hypothetical cases using a limited amount of data, resulting in a high level of uncertainty regarding the usefulness of the tool's outputs for the Greek agricultural administration. The CS experience also confirmed that uncertainties around data scalability become much greater when complementary datasets (such as structural, socio-economic or environmental data) are missing, fragmented, or are not interoperable. Furthermore, the lack of engagement resulted in a clear gap between the demonstration of the tool and the link to policy application.

The CS authors' main recommendation is that Greece should work towards robust data integration, iterative stakeholder engagement, and flexible scaling across crop and farm levels. This would mean ensuring interoperability with IACS/LPIS and FADN/FSDN systems, co-defining indicator thresholds with national authorities, and embedding workshops and questionnaires at multiple stages of tool implementation to ensure that user feedback shapes its refinement.

Case study Italy – Testing the Policy Monitoring Tool (PMT) to strengthen CAP performance evaluation

In the context of the CAP Strategic Plan 2023–2027, Italy sought tools that could simplify reporting while improving the analytical basis for evaluating the effectiveness of interventions. The Italian CS tested the Policy Monitoring Tool (PMT), which integrates multiple datasets and visualises performance indicators in a unified dashboard. CREA (Italian institute for agricultural research and agricultural economics) tested whether the PMT could support ex-post evaluation of agricultural policy by capturing key performance indicators, measuring the progress towards goals, and ensuring alignment with the country's strategic priorities. Furthermore, the results were intended to reflect the tool's ability to provide timely and actionable insights to policymakers and stakeholders and adapt to evolving policy priorities, and consequently improve accountability and transparency in implementation.

During the CS, synthetic datasets were used to demonstrate PMT's ability to generate interactive visualisations across a range of CAP indicators, including socio-economic results and environmental outcomes. Two national workshops were held with CREA to demonstrate and discuss the tool. Initially conceived mainly for ex-post evaluation, the structured (though limited) feedback collected during these sessions highlighted that the PMT could also effectively support ex-ante assessments – including scenario testing – alongside improved usability, indicator relevance, and potential integration with administrative systems. These findings helped reveal the PMT's broader value for harmonising indicator monitoring and strengthening evidence-based decision-making. Respondents provided structured feedback on the tool's functionalities, such as the need for a clearer user interface, options for bulk downloads, and the integration of farm-level data to complement crop-level analyses. Overall, the tool was seen as promising, but still under development, with the potential to become a valuable support tool for policy support if it were expanded and better integrated with existing data infrastructures.

Over the course of the CS, several challenges surfaced, especially a lack of access to national administrative data which was to be provided by CREA, and the somewhat limited response to the structured feedback questionnaire. The low response rate to feedback questionnaires perhaps indicated limited institutional capacity or willingness to invest in new digital evaluation tools.

The CS recommends that future efforts to implementation the PMT begin with a longer preparatory phase to secure access to relevant official datasets and strengthen early engagement with institutional partners to ensure policy relevance. If Italy (or any other country) were to embed the PMT within its national monitoring framework, this would require formal data-sharing agreements, improved interoperability and capacity-building across administrations. Collaboration would benefit from a broader team with greater data skills and from a gradual rollout supported by intermediate demonstrations and regular feedback loops. In addition, the tool would benefit from enhancements to make it more user-friendly, automated data-quality controls and strengthening links between the PMT's outputs and CAP performance reviews. Further research could explore the PMT's capacity for predictive analysis and scenario testing. If institutionalised under these conditions, the PMT could substantially improve transparency, timeliness and analytical depth in Italy's CAP monitoring and evaluation system.

Case study Spain (2) – SOC Map: Use of Earth Observation and Machine Learning to estimate Soil Organic Carbon (SOC)

Spain's commitments under CAP 2023–2027 and National Integrated Energy and Climate Plan highlight the need for cost-effective, scalable methods to monitor Soil Organic Carbon (SOC), a key indicator for soil-based climate mitigation. Traditional soil sampling is accurate, but it is also costly and spatially limited. The objective of this case study was to test whether combining a national database of georeferenced soil samples, Copernicus satellite data, auxiliary datasets and machine-learning algorithms can generate SOC maps that support CAP monitoring and the evaluation of carbon-farming measures.

The CS developed a national-scale SOC estimation workflow using open-source tools and a variety of algorithms (Random Forest, Cubist, Partial Least Squares Regression – PLSR). Bare-soil spectral composites, climatic variables, topography and land-use data were integrated to produce a SOC map of Spain. The prototype map demonstrated that remote sensing can significantly expand spatial coverage at relatively low cost, providing baseline

layers for CAP evaluation. The workflow included uncertainty mapping, a critical innovation for assessing the reliability of estimates and guiding future sampling design.

The findings show that the method could complement traditional surveys and reduce monitoring costs, but representativeness issues remain due to uneven soil-sampling density and vegetation cover affecting spectral data quality. Risks include overconfidence in model outputs without proper communication about uncertainty and limited institutional capacity to produce, maintain and update SOC maps, as this kind of effort requires expertise in soil science, GIS management, machine learning and data science.

The CS recommends investing in harmonised soil-sampling protocols, expanding field datasets, and building national capacity in remote sensing. SOC maps could be integrated into CAP monitoring frameworks and connected with EU-level initiatives such as the Soil Observatory. With further refinement, the method can support evidence-based design of carbon-farming schemes and provide a scalable approach for MSs seeking to monitor soil indicators efficiently.

2.3.2. TEF analysis

ACCURACY

Accuracy analysis considers how successful the tool is in offering results that reflect reality without any distortion or omission. The results of the evaluation are presented below.

DATA ACCURACY AND VALIDITY

Data accuracy and validity refers to the extent to which data input and its sources are traceable and validated.

Evaluation Question: To what extent do the tools rely on well-founded evidence (e.g. the sources of data used during the implementation of the tool are known to be reliable)? Is it possible to easily trace the information used for supporting the policy choices back to its original source? Are measures in place to avoid mistakes and redundancy and to ensure consistency in the data or information used to support the policy choices (e.g. to verify the quality and reliability of the information)? Are results triangulated, corroborated, or substantiated by other data sources?

For the monitoring-oriented tools, data accuracy and validity depend directly on the quality of underlying administrative, statistical and Earth Observation (EO) datasets. The Policy Monitoring Tool (PMT) prototypes in Italy and Greece can be linked conceptually to the CAP Performance Monitoring and Evaluation Framework (PMEF), though this linkage is currently not yet finalised. In principle, this ties the PMT outputs to official, documented and regularly updated sources such as IACS/LPIS, FADN, national statistics and environmental data. However, both of these CSs highlight that fragmented databases, weak interoperability and incomplete national registers limit validity. These interoperability challenges further amplify aggregation difficulties and undermine the accuracy of the resulting indicators. Furthermore, a significant obstacle now is the access to administrative data. If this obstacle is overcome, however, it should be possible to trace each dashboard indicator back to its original data table, with clear metadata on definitions, time coverage and units, to avoid potential misinterpretation of outputs.

The Spanish CS for the SOC mapping tool uses georeferenced soil samples combined with Copernicus imagery and other environmental data in machine-learning models to estimate Soil Organic Carbon (SOC). This approach is well supported in the digital soil-mapping literature, which shows that EO-based ML models can predict SOC when sampling is well designed and models are carefully validated (e.g. Ji et al., 2024; Yuzugullu et al., 2024). However, SOC maps remain only as accurate as the training data they are based on: sparse or biased sampling, lab inconsistencies and unbalanced land-use representation can introduce systematic errors. Residual vegetation, soil moisture and surface roughness also affect spectral signatures.

In all three monitoring cases, rigorous quality control, documentation of data provenance and triangulation with independent references (e.g. legacy soil maps, existing indicators, expert review) are essential for credible results.

TOOL REPRESENTATIVENESS

The capacity of a tool to accurately and comprehensively capture or depict a system, dataset, or phenomenon. It combines the notions of accuracy and coverage in its depiction.

Evaluation Question: To what extent do these tools exhibit a high degree of sensitivity (i.e. how much are the tools' outputs influenced by changes in the input data)? Are the tools able to capture the diverse aspects, variations, or elements within the system or dataset they seek to represent?

The representativeness of PMT depends on the breadth and quality of indicator sets that its operators can access. A well-implemented PMT could probably approximate a holistic picture of CAP performance, if based on quality national PMEF datasets in conjunction with national IACS/LPIS databases. In practice, outputs will be more representative for well-measured domains (e.g. area-based interventions, basic income support, some environmental indicators) and weaker for emerging or data-poor themes (e.g. social impacts, specific biodiversity outcomes).

The SOC mapping tool shows high sensitivity to input data, with model performance strongly affected by the number, quality and spatial distribution of soil samples – poorly sampled or unsampled areas reduce predictive reliability and increase uncertainty. Sensitivity also arises from satellite-image conditions such as cloud cover that impedes the effectiveness of bare-soil filtering, all of which introduce noise into the model and therefore, influences model accuracy. The opacity of machine-learning algorithms can make it difficult to diagnose how input variations affect outputs or to assess model behaviour in any other way than actually running the model and examining its performance metrics. In terms of representativeness, the tool can capture diverse spatial patterns when supplied with dense, well-distributed soil samples and relevant auxiliary variables (topography, climate, land use), but reliability declines where bare soil is rarely visible or historical soil data are limited or unreliable. Overall, the methodology can represent SOC variability at multiple scales, but results must be interpreted with explicit attention to data constraints, sample-density differences and model uncertainty.

RELIABILITY

The reliability assessment aims at measuring the consistency and stability of the results obtained by the application of the selected tools. This criterion measures the extent to which the process of implementation of the tools produces consistent findings when repeated. The results of the evaluation are presented below.

REPLICABILITY

Replicability refers to the extent to which the tool's results are (or can be expected to be) consistent across studies aimed at answering the same question, either by replicating the same approach or using new data or methods.

Evaluation Question: Can the results of the sub-category of tools be replicated successfully by others or through different approaches? Do the results remain consistent with other studies and approaches?

The monitoring tools are, in principle, highly replicable when the steps in the process, indicator definitions, and data-handling procedures are standardised and well documented. PMT outputs can be reproduced by other teams if they have access to equivalent datasets and full information about pre-processing choices and data-management settings. Because the PMT primarily reuses existing administrative and EO data, repeating the workflow with updated inputs should yield consistent results, provided data structures and indicator definitions remain stable. Any changes in administrative coding systems, data quality, or PMEF revisions must be carefully logged and versioned to preserve comparability across years.

For the Spanish SOC map, replicability depends on having clear, open documentation about how the data were prepared, which modelling approaches were used, and what settings (training parameters and performance metrics) guided the process. Digital soil-mapping studies emphasise cross-validation and independent test sets to check how well a model performs on new areas beyond the samples it was trained on (Ji et al., 2024). It is also important to note that different machine-learning algorithms will produce different SOC maps even when trained on the same data and with similar settings. This reflects the fact that each model learns patterns in a different way. For this reason, the choice of algorithm must be clearly documented to ensure that future users can reproduce results or understand why different versions of a map may diverge. If the Spanish workflow is shared (scripts, parameter files, training data), other teams can rerun the process in Spain or adapt it to other countries. Without this information, it becomes difficult to reproduce results or to understand why two SOC maps differ.

TRANSPARENCY

Transparency refers to the extent to which information concerning underlying methodologies, data, and results of the tool are available and accessible, as well as the extent to which the tool is open to external review.

Evaluation Question: is the information concerning underlying methodologies, data, and results of the tool available and accessible, and is the tool accessible for external review?

As with accuracy, the first issue to be resolved in these tools is the underlying datasets. Transparency in reporting methodological aspects of indicator definition, data collection and processing, reporting, etc., is the fundamental prerequisite for further applications of any of the monitoring tools covered.

In a well-designed PMT for Italy or Greece, users should be able to move easily from a graph or map to the underlying indicator definition, the data source, the period and area covered, and ideally the key processing steps. This reflects good practice in indicator-based monitoring systems such as the PMEF. In reality, however, the pilots show that transparency can be limited – often because systems are developed by small technical teams with limited documentation, or because data-protection rules restrict access. Publishing methodological summaries and metadata, even in simplified form, is therefore essential for transparency.

Producing maps such as the Spanish SOC map can be highly transparent when it relies on open satellite data (e.g. Copernicus Sentinel), open-source software and publicly available modelling procedures. Nonetheless, machine-learning methods are often seen as "black boxes" by non-specialists. Ensuring transparency therefore requires plain-language explanations of the inputs used, the main modelling steps, the validation approach and the degree of uncertainty in the results, together with clear and intuitive map legends. Recent EO and digital soil-mapping studies (Ji et al., 2024) increasingly stress the importance of communicating uncertainty and limitations alongside visually appealing maps, and the Spanish CS follows this direction.

Applicability

The applicability assessment aims at measuring the extent to which the tool can be used to different contexts and pursue different objectives. The results of the evaluation are presented below.

Multi-functionality

Multi-functionality refers to the extent to which the tool pursues several objectives and policy-making steps at the same time, (i.e., the number of policy objectives or policy-making steps that can be addressed by the tool), as well as the extent to which the underlying assumptions of the tool can be easily changed to address other policymaking steps or other objectives.

Evaluation Question: To what extent can these tools be used to address different policy objectives and steps of the policymaking process? Has the sub-category of tools been used to contribute to one or more of the CAP Specific objectives and to one or more steps of the CAP Strategic Plan design or monitoring? Can the tool be tailored or re-adjusted to assess or contribute to the assessment of the other Specific Objectives and policymaking steps?

PMT is inherently multi-functional. Once connected to relevant data sources, the two CSs have revealed that it could support ex-ante, interim and ex-post evaluations**.** Because it is indicator-driven, it can also address multiple CAP specific objectives, subject to data availability. It can hypothetically also support other policy frameworks (e.g. climate, biodiversity strategies) by reusing the same core infrastructure. Based on stakeholder interactions, a way to increase overall validity for policymaking would be to adopt a farm-based statistical approach, rather than a purely land-based approach with area-based indicators.

The SOC map is far more specialised, but may still support several policy tasks: establishing baseline SOC stocks, targeting carbon-farming schemes, evaluating the effects of management changes or climate scenarios, etc. Its applicability is strongest where soil carbon is a clearly articulated policy priority; in such cases, it could become a component rather than a standalone monitoring system. Furthermore, the underlying methodology can support a wide range of policy tasks. Because the workflow relies on broadly available 'ingredients' such as open satellite data, auxiliary environmental variables and machine-learning approaches, the approach can be adapted to different regions, sectors or environmental themes, even if the specific Spanish prototype is not directly transferable. In this sense, SOC mapping functions more as a modular component within wider monitoring and modelling systems rather than a standalone tool, and its availability depends primarily on access to data and technical capacity rather than on proprietary software.

Experience

Experience refers to the extent to which the tool has been used across MSs, in previous programming periods, or in other policies.

Evaluation Question: How experienced are MSs with this sub-category of tools? Has the tool been used in previous programming periods of the Common Agricultural Policy (CAP), or has the tool been used for other policies (e.g. climate, environmental etc.)? How many MSs have used the tool in this CAP programming period?

Performance- and indicator-based monitoring has a long history in the CAP, from the CMEF in the 2007–2013 and 2014–2020 periods to the PMEF under the 2023–2027 CSPs, and it is central to the New Delivery Model (NDM). Projects such as MEF4CAP have documented existing monitoring systems and explored pathways for integrating new data sources – such as Earth Observation, sensor data and big-data analytics – into future CAP monitoring. However, bringing these heterogeneous data sources together in a coherent operational dashboard remains a challenge. The Italian and Greek PMT pilots build on this wider body of experience, but represent early national implementations; they demonstrate technical feasibility, but still facing institutional, data-governance and capacity-related barriers to full-scale deployment.

For SOC mapping, there is now substantial international experience combining different data sources with machine-learning approaches to map SOC and soil organic matter at multiple scales (Zhang et al., 2017; Ji et al., 2024; Yuzugullu et al., 2024). However, operational integration of such products into formal CAP monitoring and reporting is still emerging. The Spanish SOC map therefore stands at the interface between research and policy practice: scientifically mature methods are being developed and tested in a governance context, but the adjustment of these methods to digital and EO-based monitoring has yet to be achieved. It is likely that integration into 'regular' monitoring end evaluation procedures will require a substantial period of testing, administrative capacity-building and integration with existing systems.

Granularity

Granularity refers to the extent to which the tool can provide results at different level of spatial or temporal resolution (e.g. local, regional, national, and EU level).

Evaluation Question: Can the tools support analysis and policymaking at different levels, including local, regional, national and EU level? Do the tools provide insights at different geographical and time scales (e.g. results expressed in days, years, weeks)?

The granularity of PMT is entirely dependent on the input data. In principle, it can provide outputs at any spatial scale ranging from land parcel level to the EU in total. Temporal granularity is currently typically annual, aligned with CAP reporting cycles, though underlying data sources (e.g. EO-derived crop maps, payment data streams) can support more frequent analysis. This flexibility makes PMT suitable for both high-level dashboards and more detailed internal diagnostics.

SOC map offers very high spatial granularity (pixel- or field-scale SOC estimates at ~10–20 m) but rather coarse temporal granularity. SOC changes slowly, so meaningful updates are usually planned at multi-year intervals (e.g. 5–10 years). The approach can thus inform medium- to long-term assessments and target setting, but is less suited to annual performance tracking due to the relatively slow overall changes brought about by soil management interventions.

Accessibility

The accessibility assessment aims to evaluate the extent to which the tools and their underlying methodologies, data, assumptions, and outputs are clear and accessible to stakeholders and users. The results of the evaluation are presented below.

Ease of use

Ease of use refers to the extent to which the tool is easy to learn and use.

Evaluation Question: Does the subcategory of tools require prior training or experience, or does it require much time to be able to use it properly? Can these tools be handled by a wide range of professionals, or do they require specific professional profiles to be implemented?

For end-users, PMT dashboards are conceptually similar to common business-intelligence tools: once built, they are relatively easy to navigate, filter and interpret, assuming adequate training on indicator definitions and caveats. The main complexity lies acquiring data, managing databases and ensuring data security and interoperability; this

requires specialised IT and data-engineering skills that are not necessarily widely available in all CAP administrations.

The SOC map is technically demanding to produce. It requires EO processing, GIS and machine-learning expertise, but should be relatively straightforward to use and interpret once maps are generated; however, this assumption is still to be corroborated in the Spanish CS. In principle, distribution via web viewers or GIS services should allow policymakers, advisors and even farmers to consult changes in SOC visually. The key usability challenge lies in communicating uncertainty, especially in areas with sparse training data, and in avoiding over-interpretation of pixel-level variation.

Tool Availability

Availability refers to the extent to which the tool is accessible, available to and usable by end users.

Evaluation Question: Are these tools openly available to everyone, and their methodological and technical guidelines accessible to users? What are the main factors the limit its availability (e.g. costs, language, ease of use, available support, complexity, technical requirements)? Are there legal or patent-related limitations to the use?

The source code of PMT is openly available. Many software options (open-source and commercial) can support PMT-like functionality; the main constraints are access to data (including various data privacy issues and interoperability barriers) and institutional willingness to invest in integration. Methodological material on indicator-based monitoring and PMEF implementation is also public.

The basis for the SOC mapping approach was likewise intentionally built on widely available components: Copernicus Sentinel data (free and open), open-source GIS and ML libraries, and national soil data. Limiting factors in other MS/regions aiming to replicate the approach may be fragmented or non-open soil datasets, lack of computing resources and scarcity of in-house EO/modelling expertise in administrations. Where data and skills exist, the methods can be transferred and adapted; where they do not, MSs may depend on engaging external research capacity and/or EU-level platforms for data.

Efficiency

The efficiency assessment aims to evaluate the extent to which the tools are able to achieve desired outcomes given the resources available, including time, human resources, and costs. The results of the evaluation are presented below.

Time Efficiency

Time efficiency refers to the extent to which the tool is able to achieve its outcomes within a reasonable amount of time. It includes the time necessary for its set up and for the implementation.

Evaluation Question: How much time does it take to set up this group of tools (i.e. time necessary to organise and prepare the implementation)? And how much time does it take to implement the tool until the achievement of the intended outputs?

The Greek and Italian pilots demonstrated the conceptual feasibility of the PMT rather than its operational efficiency. While a fully developed PMT could, in principle, automate updates once stable data pipelines and indicator definitions are in place, the pilots did not reach this stage. Instead, they highlighted the substantial preparatory effort required to secure datasets, document data structures, and coordinate with institutions—steps that must be completed before any automation or routine reporting becomes possible. It can reasonably be expected that further work to resolve interoperability issues, harmonise data formats, and address data-quality and data-management challenges would take from several weeks to several months, depending on the volume and heterogeneity of the data involved, and complexity of issues tackled.

SOC mapping also requires a substantial initial phase: assembling and harmonising soil datasets, designing or reviewing sampling strategies (where new samples are needed), resolving methodological issues such as testing alternative modelling approaches. Once this groundwork is completed, producing updated SOC layers – whether for additional years, larger regions or refined resolutions – becomes considerably faster, especially if satellite imagery, pre-processing routines and code are systematically archived. Over the medium term, EO-based SOC estimation can be more time-efficient than repeated dense field-sampling campaigns, particularly for larger territories where traditional sampling is costly and logistically demanding. This tool also can be used to identify those less-sampled areas on which focus sampling campaigns.

Cost-effectiveness

Cost-effectiveness refers to the extent to which the tool is able to achieve its outcomes within a reasonable budget. It includes the budget necessary for patent subscriptions, maintenance, and implementation.

Evaluation Question: Is implementing these tools expensive? How much budget is needed for their implementation? Is acquiring and/or maintaining them expensive? How much budget is required for the maintenance of the tool or subscription for patent-based tools every year?

PMT leverages existing data (LPIS/IACS, statistics, survey data) and repurposes them, potentially for multiple needs – CAP reporting, national strategy monitoring, other EU obligations. By reducing manual data handling, duplication and ad-hoc analysis and reporting, PMT can free up staff time.

SOC mapping requires funding for soil sampling, lab analysis, EO processing and modelling. However, studies increasingly underline that remote-sensing-supported SOC mapping can reduce monitoring costs per hectare compared to dense traditional sampling, especially when EO infrastructure and models are reused for other soil variables or applications (e.g. clay, texture, erosion) (Ji et al., 2024). Cost-effectiveness may further improve when such mapping is integrated into broader soil-health and climate-monitoring strategies rather than treated as a one-off project.

Human Resources Efficiency

Human resource efficiency refers to the extent to which the tool is able to achieve its outcomes involving a reasonable number of human resources.

Evaluation Question: How many people are necessary for the implementation of the tool and the analysis of results? How many different professionality are necessary for the good implementation of the tool?

Both PMT and SOC mapping concentrate technical workload in small specialised teams while serving a broad audience of users. For PMT, a handful of data engineers and analysts can maintain the system, while many policy officers, evaluators and external stakeholders benefit from the outputs with minimal additional HR cost. For SOC mapping, a relatively small EO/modelling team can produce spatial products that inform multiple instruments (eco-schemes, AECM, carbon-farming pilots, advisory services).

In summary, the Italian, Greek and Spanish monitoring case studies show that digital monitoring tools and EO-based products can significantly strengthen evidence-based CAP implementation, provided that Member States invest in data infrastructure, maintain clear governance and intentionally design user-friendly outputs. They are not low-effort solutions, but when embedded into long-term strategies, they offer a favourable balance between analytical power and resource requirements.

2.3.3. Key findings

This section summarises key aspects of the monitoring tools evaluated in Italy, Greece and Spain. These tools are designed to support the CAP's performance-based governance by integrating administrative, environmental and Earth-Observation data from various sources to enable systematic tracking of policy results and environmental outcomes.

Based on the evaluation, the main results can be summarised as follows:

Accuracy: PMT accuracy depends on the completeness, consistency and interoperability of underlying administrative datasets. Fragmented national data systems will therefore reduce accuracy. SOC Map accuracy depends on soil sampling quality, spectral signal interpretation and proper model validation. Both tools improve in accuracy when data pipelines, metadata and QA/QC (Quality assurance/Quality Control) procedures are well established.

Reliability: PMT workflows can be repeated consistently when indicator definitions remain stable. SOC Map is reliable if modelling workflows are documented and reproducible. Both tools are sensitive to updates in baseline data or sampling regimes but can yield consistent trends and patterns when version control and documentation are applied. Reliability increases with stable governance and consistent data management.

Applicability: PMT can be used for ex-ante baselines, annual (interim) monitoring and ex-post evaluation, making it broadly useful across CAP policy cycles. SOC Map can support soil-health assessments, carbon-farming targeting

and climate reporting. However, applicability is limited where data interoperability, EO capacity or soil datasets are incomplete. Both tools require adaptation before being transferred to other Member States.

Accessibility: While dashboards and maps are easy for policymakers to interpret, the production of these tools requires specialised expertise – data engineering (PMT), GIS and machine learning (SOC Map). Their implementation depends on the availability of technical staff, stable IT infrastructure and institutional cooperation. Accessibility improves when workflows are open-source and documentation is complete but is significantly lower compared to qualitative tools.

Efficiency: Both PMT and SOC Map require significant upfront investment: system integration, indicator harmonisation, sampling campaigns, EO processing and modelling – even when the input data already exist; additional costs may be associated with indicator definition and data collection. They offer efficiency gains only after initial development – automation of data analysis and reporting (PMT) and reduced long-term soil-monitoring costs (SOC Map). Short-term efficiency is low, but long-term efficiency can be high when systems are maintained and reused across policy cycles.

Monitoring tools offer substantial long-term value for evidence-based CAP implementation. They enable structured, scalable and spatially rich monitoring systems aligned with the New Delivery Model. Their main limitations concern technical capacity requirements, potentially high initial costs, data interoperability challenges and dependence on specialised expertise. Sustainable impact requires investment in data infrastructure, institutional coordination and continuous methodological development.

2.4. Integrated cases

2.4.1. Summary of the Case study reports

Case study Netherlands – Integrated modelling and monitoring framework for eco-scheme design and evaluation

The Netherlands is required to align its highly intensive agricultural system characterised by high productivity, a strong market orientation and substantial environmental pressures (nutrients, biodiversity decline, GHG emissions) with the EU's environmental ambitions. The Dutch CAP Strategic Plan (2023–2027) places considerable emphasis on eco-schemes, for which robust design and calibration are crucial to ensure their effective operation. In the Dutch eco-scheme implementation, farmers collect points for implementing certain activities (e.g., permanent grassland, landscape features), which determine an associated payment class with three tiers: bronze (€60/ha), silver (€100/ha), gold (€200/ha). In 2023, there were 35,000 applications under the eco-scheme (70% of all farms), which exceeded the available budget.

The aim of the Dutch CS was to assess how an integrated toolbox – linking qualitative, quantitative and monitoring tools – can support evidence-based eco-scheme design. The integrated approach combined: (i) an Intervention–Objective Interaction (IOI) matrix to examine policy coherence; (ii) the FARMDYN[6] farm-level optimisation model and the AGMEMOD[7] market model to quantify economic and environmental impacts of alternative eco-scheme designs; and (iii) a Policy Monitoring Tool (PMT) exploring options for digital compliance and performance verification. This allowed policymakers to explore synergies and trade-offs across objectives, behavioural reactions to eco-scheme incentives, and the feasibility of monitoring through administrative and geospatial data.

Methodologically speaking, this case study required significant integration across models and data systems. The IOI matrix was used to identify relevant eco-scheme attributes and potential inconsistencies with other CAP interventions. FARMDYN modelled changes in farm behaviour, production structure, and income under alternative payments and related requirements, while AGMEMOD complemented this by capturing market-level adjustments. The PMT component assessed how the Area Monitoring System (AMS), the Land Parcel Identification System (LPIS) and other geospatial layers could be used for automated verification of eco-scheme practices. The results highlight that eco-scheme effectiveness depends strongly on payment calibration, regional heterogeneity, and the behavioural response of specialised dairy farms. Some eco-scheme options produced meaningful environmental improvements, but involved higher compliance costs or required more advanced monitoring infrastructure. The IOI

[6] https://www.ilr1.uni-bonn.de/en/research/research-groups/economic-modeling-of-agricultural-systems/farmdyn

[7] https://agmemod.eu/

results showed several coherence gaps between Pillar I eco-schemes and the Pillar II Agri Environmental Climate Measures (AECM), reinforcing the need for integrated planning across the two Pillars.

The CS demonstrates the potential for combining tools to analyse the specific policy mix in CAP strategic planning, but also reveals institutional challenges. The time and expertise required to integrate modelling tools with monitoring systems were underestimated, especially given the need to harmonise data sources and ensure model compatibility. Effective use of such an integrated toolbox requires a stable partnership between modelling institutions, IT/monitoring units, and ministerial policy departments, with long-term investments in data infrastructure and modelling capacity required. The experience suggests that integrated modelling–monitoring frameworks can significantly improve CAP design, but only if supported by permanent institutional collaboration and continued development of national datasets.

Case study Slovenia: Integrated Tool for Agricultural Policy Vision Building in Slovenia

Slovenia faces the dual challenge of ensuring the long-term viability of its small-scale, diversified agricultural structure, while responding to EU priorities such as climate adaptation, biodiversity protection and rural well-being. To address these challenges, the Slovenian Ministry of Agriculture developed a comprehensive Vision for Agriculture and Food 2040 (Vision 2040), supported by an integrated toolbox combining strategic foresight, scenario analysis, and indicator-based monitoring. The development and direct application of the integrated toolbox in Vision 2040 was the subject of the Slovenian CS.

The objective was to strengthen strategic planning using participatory, data-driven tools that increase policy legitimacy and transparency. The research group aimed to demonstrate how combining participatory tools, quantitative indicators and trend analysis can support long-term strategic planning and guide future CAP programming. The approach included development of the methodology, desk research, discussion of evidence within expert panels, stakeholder consultations, and a set of interactive scenario workshops that explored priorities, mechanisms and alternative futures. These were linked with a large set of sustainability indicators derived from national datasets and the PMEF framework.

Methodologically speaking, the CS required the development and harmonisation of more than 90 indicators, based on the CSP context indicators, spanning resources, production, income, productivity, climate impact and resilience, environmental quality, biodiversity, generational renewal and rural quality of life. Trends were analysed and discussed to identify an ideal (while still realistic) value of selected indicators in 2040. The scenario work clarified where Slovenian agriculture may face persistent pressures (e.g., abandonment in marginal areas, structural income gaps, climatic impacts on grasslands), and where policy interventions are particularly needed to achieve long-term goals. The integrated framework highlighted trade-offs (and potential synergies) between environmental goals and farm income, the centrality of knowledge and innovation systems, and the importance of structural change. The monitoring component showed how annual and multi-year indicators could be used to track progress toward 2040 objectives and inform future CSPs.

The case study shows that integrated strategic foresight and monitoring can substantially strengthen long-term policymaking but require institutional commitment and continuous research capacity. The development of indicators, stakeholder processes and scenario work were more time- and resource-intensive than expected, and data gaps limited the robustness of some of the sustainability dimensions. Maintaining the monitoring system and ensuring it actively informs policymaking into the future will require dedicated staff, cross-ministerial cooperation and stable political support. As was found in the Dutch CS, the Slovenian CS suggests that integrated approaches are powerful, but require long-term investment, iterative refinement, and sustained engagement from both policymakers and researchers. Graphical representations of trends and the desired scenario were key to communicating with stakeholders.

The greatest added value in the integrated toolbox that has been developed is a structured approach to discussion and linking policy-relevant data with the vision process. Limitations were evident due to missing data (social dimension, value chain). The provision of support for policymakers via the integrated toolbox is an opportunity to strengthen institutional capacities and contributes to greater transparency in policy-making, but it has also revealed data gaps and the importance of addressing them to strengthen evidence-based policy-making. The toolbox has been tested in the real world of policy design and forms the basis for the CAP after 2028. The Slovenian CS shows that this approach can be used and integrated into routine strategic planning processes in the field of agricultural policy, and that the tool is transferable. In the future, more effort needs to be focused on improving the database (especially value chain and social indicators). The development of more advanced scenario methods (e.g., system dynamics) should also be considered, provided sufficient resources are available.

2.4.2. TEF analysis

ACCURACY

Accuracy analysis considers how successful is the tool in offering results that reflect reality without any distortion or omission in those results. The results of the evaluation are presented below.

DATA ACCURACY AND VALIDITY

Data accuracy and validity refers to the extent to which data input and its sources are traceable and validated.

Evaluation Question: To what extent do the integrated toolboxes rely on well-founded evidence? Is it possible to easily trace the information used for supporting the policy choices back to its original source? Are measures in place to avoid mistakes and redundancy and to ensure consistency in the data or information used to support the policy choices (e.g. to verify the quality and reliability of the information)? Are results triangulated, corroborated, or substantiated by other data sources?

Based on the Dutch CS, the integrated toolbox draws heavily on FARMDYN and AGMEMOD, both of which rely on carefully curated datasets and assumptions documented in peer-reviewed publications (Britz et al., 2021; Chantreuil et al., 2012). These models use official statistical sources such as Eurostat, FADN, and national economic accounts, which improves traceability and internal consistency. The underlying theoretical basis of micro-economic programming models (FARMDYN) ensures that behavioural responses are grounded in profit-maximisation principles (Hazell & Norton, 1986), while partial-equilibrium models, such as AGMEMOD, follow standard market-modelling frameworks. In addition, the monitoring component of the Dutch toolbox, based on LPIS/IACS and the PMEF logic, is grounded in administrative data that are routinely validated through EU audit processes (European Commission, 2022).

Based on the Slovenian CS, data accuracy rests on indicator frameworks that use official national statistics, Eurostat, environmental datasets, and CAP monitoring elements. Indicator-based approaches align with international guidance on sustainability assessment (FAO, 2021; OECD, 2017). Foresight and scenario-building processes rely on expert knowledge, which is consistent with standard strategic foresight practice (OECD, 2019; UNDP, 2018), but is inherently less "empirically verifiable."

Across both of these integration-based CSs, the integration across multiple tools increases data robustness through triangulation, an approach well established in evaluation theory (Patton, 2015). However, limitations arise from the need to harmonise datasets across models and the incomplete coverage of certain environmental or social indicators, consistent with gaps noted in EU monitoring literature (MEF4CAP, 2023; Hargreave et al., 2022).

TOOL REPRESENTATIVENESS

The capacity of a tool to accurately and comprehensively capture or depict a system, dataset, or phenomenon. It combines the notions of accuracy and coverage in its depiction.

Evaluation Question: To what extent do these two integrated toolboxes present a high degree of sensitivity (i.e. how much the results of the tools are influenced by changes in the input data)? Are the toolboxes able to capture the diverse facets, variations, or elements within the system or dataset they seek to represent?

The Dutch toolbox combines: i) farm-level modelling (FARMDYN) capturing farm heterogeneity and technology choices (Britz et al., 2021), ii) market-level modelling (AGMEMOD) capturing trade and price adjustments (Chantreuil et al., 2012), iii) policy coherence tools (IOI matrix), and iv) monitoring indicators linked to spatial data (PMEF, LPIS). This multi-scale structure increases representativeness and aligns with best practice in integrated assessment modelling (van Ittersum et al., 2008; Gonzalez-Martinez et al, 2021). Such representativeness is sufficient for the purpose for which the toolbox was selected, namely to provide a comprehensive overview and assessment of eco-scheme interventions.

The Slovenian toolbox emphasises thematic/strategic representativeness, capturing economic, social, and environmental dimensions of sustainability through around 100 indicators aligned with the EU PMEF. The proposed participatory approach to supporting vision development both utilised and reinforced the work with hard data, strengthening representativeness.

RELIABILITY

The reliability assessment aims at measuring the consistency and stability of the results obtained in the application of the selected tools. This criterion measures the extent to which the process of implementation of the tools produces consistent findings when repeated. The results of the evaluation are presented below.

REPLICABILITY

Replicability refers to the extent to which the tool's results are (or can be expected to be) consistent across studies aimed at answering the same question, either by replicating the same approach or using new data or methods.

Evaluation Question: Can the results of the integrated toolboxes be replicated successfully by others or through different approaches? Do the results remain consistent with other studies and approaches?

In the Dutch CS, replicability is the highest for the two models included, namely FARMDYN and AGMEMOD, both of which are part of established modelling communities and documented in scientific literature (Britz et al., 2021; van Leeuwen et al., 2020). Under identical assumptions, simulation models generally produce stable results (Petersen et al., 2018). However, integrated runs that combine outputs across models inevitably introduce coordination challenges and scenario design differences that hinder exact reproducibility, a well-known issue in integrated assessment modelling (Kelly et al., 2013). When designing the IOI Matrix, which involves multiple checks of qualitative assessments of potential effects and coherence, it is possible to ensure replicability to a certain extent, but this is of course not absolute. The results of the IOI Matrix also help to verify the results of both models, which points to the need for a broader analytical approach with more powerful tools (such as FARMDYN and AGMEMOD) when dealing with such a complex issue as the economic and environmental effects of eco-schemes.

In the Slovenian CS, replicability is strong for the indicator-based monitoring, as indicator definitions and extraction procedures can be standardised. Foresight elements, however, are inherently non-replicable in a strict sense, as results depend on more subjective factors relating to stakeholder participation, expert judgement and evolving contexts (UNDP, 2018; OECD, 2019). This aligns with foresight practice, where validity is achieved through adhering to a functional process rather than assessing predictive alignment (Miles, 2010).

TRANSPARENCY

Transparency refers to the extent to which information concerning underlying methodologies, data, and results of the tool are available and accessible, as well as the extent to which the tool is open to external review.

Evaluation Question: is the information concerning underlying methodologies, data, and results of the integrated toolboxes available and accessible, and is the toolbox itself accessible for external review?

In the Netherlands, the underlying models (FARMDYN, AGMEMOD) are publicly accessible with existing theoretical documentation and peer-reviewed publications, but they are complex and require specialised expertise to interpret (Britz et al., 2021; Chantreuil et al., 2012). Transparency is high at the conceptual level, but partial at the level of full model code or parameterisation – mirroring the broader experience with agricultural models (Banse et al., 2015; Erjavec et al., 2018). Monitoring indicators, by contrast, have high transparency because they follow the PMEF documentation.

In Slovenia, transparency is high due to the publication of the Vision 2040 framework, indicator definitions and methodological notes. Foresight methods are inherently transparent conceptually, scenario narratives and methodological choices can be documented (van Notten, 2005), but are less transparent in terms of how expert judgement is formed.

APPLICABILITY

The applicability assessment aims at measuring the extent to which the tool can be used to different contexts and pursue different objectives. The results of the evaluation are presented below.

MULTI-FUNCTIONALITY

Multi-functionality refers to the extent to which the tool pursues several objectives and policy-making steps at the same time, (i.e., the number of policy objectives or policy-making steps that can be addressed by the tool), as well as the extent to which the underlying assumptions of the tool can be easily changed to address other policymaking steps or other objectives.

Evaluation Question: To what extent can these toolboxes be used to address different policy objectives and steps of the policymaking process? Has the toolbox been used to contribute to one or more of the CAP Specific objectives and to one or more steps of the CAP Strategic Plan design or monitoring? Can the toolbox be tailored or re-adjusted to assess or contribute to the assessment of the other Specific Objectives and policymaking steps?

The Dutch toolbox, which deals with the assessment of one set of interventions, namely eco-schemes, shows a strong and diverse level of multi-functionality: ex-ante policy assessment, farm and market impact analysis, coherence evaluation, and options for digital monitoring. This mirrors the best international practice in combining models and monitoring systems (van Ittersum et al., 2008; Erjavec et al., 2018). The toolbox addresses climate and environmental objectives in relation to economic issues. This approach can also be applied to most other CAP objectives that are of an economic and environmental nature (excluding biodiversity and territorial and social issues of rural development). The IOI matrix can be used independently to examine any objective from the set of agricultural policy concerns.

The Slovenian toolbox excels in long-term planning and integration across vision, scenarios, and indicators, consistent with foresight literature (OECD, 2019; UNDP, 2018). It supports needs assessment, strategic orientation, and future CSP alignment. The toolbox covers the entire spectrum of agricultural policy issues through Vision participatory building. However, due to insufficient data, some important agricultural policy issues are not covered by indicators (e.g., agri-food chain, nutrition, and even some fundamental issues such as the economic situation of individual groups of agricultural holdings). In any case, this is a toolbox that, by its nature, addresses the entire spectrum of agricultural policy issues.

However, applicability requires high institutional capacity. The theoretical literature confirms that integrated modelling and foresight require the maintenance of stable expert teams, long-term data investment, and strong political support (Kelly et al., 2013; Miles, 2010).

Experience

Experience refers to the extent to which the tool has been used across MSs, in previous programming periods, or in other policies.

Evaluation Question: How experienced are MSs in the implementation of the integrated toolbox? Has the toolbox been used in previous programming periods of the CAP?

Both the Dutch and Slovenian research teams have many years of experience in providing scientific and expert support to agricultural policy-makers. Wageningen University and Research (WUR) has a large team that covers various agricultural policy issues, develops quantitative tools, especially simulation models, deals with data sources, and prepares studies for the Dutch government. The Slovenian research group at the University of Ljubljana also has a history of continuous cooperation and support for agricultural policy, ranging from assistance in the preparation of strategic documents, impact assessments of various policy options, expert studies, and direct dialogue with decision-makers. Both teams have also collaborated in the past on the preparation of the CAP strategic plan in their respective MSs. Among other things, Dutch researchers have also been involved in the discussion of eco-schemes in the past, while the Slovenian team mainly worked on policy evaluation.

Both the Dutch and Slovenian CSs illustrate that integrated toolboxes require good cooperation with the ministry of agriculture and substantial expertise: modelling (NL), data engineering (NL), indicator management (SI) and stakeholder facilitation (SI). This is consistent with literature on the operationalisation of complex policy models and foresight systems (Hazell & Norton, 1986; UNDP, 2018). While outputs (dashboards, scenario narratives, indicator tables) are accessible to decision-makers, their production requires high technical capability, reducing wider accessibility.

Granularity

Granularity refers to the extent to which the tool can provide results at different level of spatial resolution (e.g. local, regional, national, and EU level).

Evaluation Question: Can the toolbox support analysis and policymaking at different levels, including local, regional, national and EU level? Do the toolbox provide insights at different geographical and time scales (e.g. results expressed in days, years, weeks)?

Both toolboxes were implemented at the MS level. They are tailored to the requirements of the legislative and policy frameworks and have clear end-users who directed the research support. Similar approaches can also be implemented in other MSs, but due to the framework set by the CAP, this is less relevant at the lower (local) or higher (EU) levels.

As noted in the modelling literature (Banse et al., 2015), integrated tools require high upfront investment, model calibration, data harmonisation, indicator development, and scenario workshops. Both toolboxes from the Netherlands and Slovenia also incorporate time-based simulation in their results.

However, long-term literature on integrated assessment and monitoring systems shows that such tools become progressively more efficient over multiple policy cycles, as infrastructures can be reused and repurposed, reducing marginal costs (Kelly et al., 2013; MEF4CAP, 2023). The development and use of integrated tools require the presence of a permanent and competent research group covering various areas of agricultural policy at the MS level and, of course, a state administration that commissions and uses the results of such work.

ACCESSIBILITY

The accessibility assessment aims to evaluate the extent to which the tools and their underlying methodologies, data, assumptions, and outputs are clear and accessible to stakeholders and users. The results of the evaluation are presented below.

EASE OF USE

Ease of use refers to the extent to which the tool is easy to learn and use.

Evaluation Question: Does the toolboxes require prior training or experience, or does it require much time to be able to use it properly? Can these toolboxes be handled by a wide range of professionals, or do they require specific professional profiles to be implemented?

Both integral toolboxes require considerable expertise to develop and operate. In the Netherlands, FARMDYN and AGMEMOD demand advanced skills in programming, econometrics and agricultural economics, similar to other macro- and meso-models used for CAP analysis. Monitoring dashboards are easier to use, but still need data-engineering support.

In Slovenia, forecasting and scenario analysis methods require specialised mediation and analytical skills, while indicator systems require staff familiar with statistics and monitoring frameworks such as the CMEF/PMEF and MEF4CAP recommendations. The results (vision documents, indicator tables) are easier for a non-technical audience to understand than the methods to acquire them. In any case, the Slovenian toolbox does not require as much highly specific knowledge and skills compared to the modelling activity within the Dutch toolbox. In the Slovenian case, experience in the field is required; on the other hand, such work does not require several years of training and model development, as is the case when using models such as AGMEMOD and FARMDYN. However, the implementation of research support for vision building, scenario analysis, and policy indicators requires a broad knowledge of various agricultural policy issues and the ability to lead participatory approaches. Something similar also applies to the IOI matrix, part of the Dutch toolbox.

We can conclude that comprehensive case studies demonstrate the importance of specific knowledge and a deep understanding of agricultural policy issues and logic. These are the conditions under which the research groups involved can support or comprehensively address individual issues (eco-schemes: the case of The Netherlands) or develop an innovative approach and support in individual phases of strategic planning (vision building) and are directly involved in actual strategic planning (the case of Slovenia).

TOOL AVAILABILITY

Availability refers to the extent to which the tool is accessible, available to and usable by end users.

Evaluation Question: Are these toolboxes openly available to everyone, and are their methodological and technical guidelines accessible to users? What are the main factors that limit their availability (e.g. costs, language, ease of use, available support, complexity, technical requirements)? Are there legal or patent-related limitations to their use?

Most components of these toolboxes are not off-the-shelf products, but belong to research networks or national administrations:

- FARMDYN and AGMEMOD are part of active modelling communities; access usually involves collaboration or agreements with others in that modelling community.
- Monitoring tools and indicator systems depend on national administrative data and IT systems – these are not normally fully open systems.
- Foresight and vision-building methods are publicly documented conceptually, but are not "packaged" in the manner that software-based modelling tools can be.

Availability is therefore shaped by institutional capacity, data governance and membership in modelling or monitoring networks, rather than licenses alone.

Efficiency

The efficiency assessment aims to evaluate the extent to which the tools are able to achieve desired outcomes given the resources available, including time, human resources, and costs. The results of the evaluation are presented below.

Time Efficiency

Time efficiency refers to the extent to which the tool can achieve its outcomes within a reasonable amount of time. It includes the time necessary for its set up and for the implementation.

Evaluation Question: How much time does it take to set up these toolboxes (i.e. time necessary to organise and prepare the implementation)? And how much time does it take to implement the tool until the achievement of the intended outputs?

As noted in the modelling literature (Banse et al., 2015), integrated tools require high upfront investment – model calibration, data harmonisation, indicator development, scenario workshops. The Netherlands faced time-intensive model integration challenges, while Slovenia required considerable stakeholder engagement and indicator-building time. The literature on integrated assessment and monitoring systems shows that such tools become more efficient over multiple policy cycles, as infrastructures can be reused, reducing marginal costs (Kelly et al., 2013; MEF4CAP, 2023).

The CSs have shown that a multifunctional and integrated approach requires a considerable investment of time and effort. In the Dutch case, this mainly involves the use of models (adaptation and validation). In the Slovenian case, it involves the amount of work required to prepare materials, include indicators, and, above all, adopt a participatory approach. In both cases, the process took 12 months, which essentially exceeded the scope of the CSs. In The Netherlands, the policy indicators were not fully implemented, while in Slovenia, after the completion of the CS (September 30, 2025), the work continued and had not been completed by the end of 2025 (i.e., the text had not yet been submitted to the government for approval). However, on November 15, 2025, the research support, which lasted exactly 12 months, from the first interviews and the development of the methodology to the preparation of the draft final texts and the scenario analysis of the indicators, was completed.

Cost-effectiveness

Cost-effectiveness refers to the extent to which the tool is able to achieve its outcomes within a reasonable budget. It includes the budget necessary for patent subscriptions, maintenance, and implementation.

Evaluation Question: Is implementing these toolboxes expensive? What budget is required for the implementation of the toolbox? Is acquiring and/or maintaining these tools expensive?

The initial set-up costs (modelling infrastructure, IT systems, stakeholder processes) are substantial, similar to other advanced policy analysis tools. However, by providing richer insights, helping to avoid poorly targeted policies and enabling re-use across multiple policy questions and programming periods, integrated toolboxes can be cost-effective in the medium to long term. They also align with EU recommendations to exploit data and models more systematically in CAP monitoring and evaluation.

From a financial perspective, based on the conducted case studies, we conclude that implementing such an integrated approach requires at least 500 to 1,000 effective hours of work (0.4–0.7 FTE) and associated costs, if no special investment in tool development is required. This amounts to between €30,000 and €70,000 in total costs

(depending on the wage level in the MS), and of course also on the agreed scope of work, the choice of tools, and their complexity. This does not include the costs at the client level in the case of participatory workshops. Here, the cost can also increase significantly, up to double the amount presented below for research support.

Human Resources Efficiency

Human resource efficiency refers to the extent to which the tool is able to achieve its outcomes involving a reasonable number of human resources.

Evaluation Question: How many people are necessary for the implementation of the toolbox and the analysis of results? How many different professions are necessary for the successful implementation of the tool?

Both cases require multi-disciplinary teams:

- The Netherlands: modellers, policy experts, statisticians and policy officers;
- Slovenia: policy experts, foresight experts, facilitators and statisticians.

This is similar to the broader experience with modelling research networks and advanced monitoring frameworks, which always rely on small but highly skilled teams. Once these teams are in place, outputs can serve a wide range of internal and external users, so human-resource efficiency increases over time, but entry costs (skills, training, coordination) are high.

The CSs have shown that broad teams are needed to provide comprehensive support with agricultural policy tools. In the Dutch case, at least five people were involved, while in the Slovenian case there were only three, but they collaborated with a wide range of experts from research and government institutions. It is difficult to determine the exact number of people needed, as this depends on the case, tools integrated, customer requirements, and the extent of the pre-existing relationship between the researchers and the ministry, but it is certainly necessary to have several people with different skillsets. For complex issues such as the evaluation of eco-schemes in the Netherlands and support for the development of the Vision in Slovenia, it is important to collaborate with different profiles of representatives from research, business, government, and non-governmental organizations.

2.4.3. Key findings

This report evaluates two CSs applied in The Netherlands and Slovenia using integrated toolboxes to support strategic planning under the CAP. The Dutch case focuses on eco-scheme impact assessment, while the Slovenian case presents a toolbox supporting strategic foresight and vision building. Integrated toolboxes combine several methods – quantitative models, monitoring systems, foresight and participatory approaches, into a coherent analytical framework.

Accuracy: The Dutch toolbox uses robust, traceable datasets and peer-reviewed models (FARMDYN, AGMEMOD). The Slovenian toolbox relies on indicators and expert-based foresight, aligned with international standards, but less verifiable empirically. In the Dutch case, accuracy is supported by the above-mentioned well-established models and official administrative and statistical data, which allow detailed quantification of economic and environmental impacts of eco-schemes. In Slovenia, accuracy rests on extensive indicator sets and documented data sources, combined with expert-based foresight. In both cases, integration across tools improves accuracy through triangulation, but also introduces new risks: results are sensitive to harmonisation of baselines, parameter choices and the coverage of indicators, especially for social and some environmental dimensions (e.g. biodiversity) where data remain weaker.

Reliability: The Dutch toolbox offers high replicability and transparency due to strong documentation, peer-reviewed papers and stable results of both models included. Transparency is limited for full model access and code. The Slovenian toolbox provides thematic breadth of the entire spectrum of sustainability and scenario-based foresight for long-term strategic planning of agricultural policy. The foresight approach in Slovenia is non-replicable in the sense that modelling can be replicated, but it is valid for strategic planning. Integrated analyses that combine several tools depend on the coordination of assumptions, while foresight and vision-building exercises in Slovenia are inherently context and participant dependent. Thus, the direction of results and the key messages are reliable, but exact quantitative outcomes may change across iterations.

Applicability: Both toolboxes support multiple CAP objectives. In The Netherlands, the integrated toolbox is capable of supporting multiple CAP requirements: eco-scheme design and calibration, trade-off analysis, coherence checks, and the design of monitoring and verification strategies. The applicability of the Slovenian toolbox is assessed as

high, but conditional on institutional capacity. Both toolboxes are conceptually transferable to other MSs, but practical applicability depends on data availability, modelling experience, and the ability to coordinate across research institutions, governmental departments and stakeholder groups.

Accessibility: Both toolboxes require specialised expertise and institutional capacity; components are not off-the-shelf. The results (dashboards, maps, narratives, indicator tables) are relatively easy to communicate to policymakers and stakeholders, but the production of integrated analyses requires high-level expertise in modelling, data engineering, statistics and foresight. Entry barriers are therefore higher than for standalone tools, and many administrations would need external support or long-term capacity building to use such toolboxes effectively.

Efficiency: The implementation of similar integrated toolboxes can take around 12 months, with high upfront costs (€30,000–€70,000) and multi-disciplinary teams. Efficiency improves over policy cycles through infrastructure reuse. Setting up integrated toolboxes is labour- and time-intensive: data harmonisation, model linkage, indicator development and participatory processes require substantial upfront investment. Once established, however, the infrastructure and methods can be reused across multiple CAP cycles and strategic exercises. This can reduce the marginal costs of additional analyses, support more coherent policy choices and lower the risk of poorly targeted or inconsistent interventions.

Both CSs show that integrated toolboxes demand long-term investment, expert teams, and institutional support, but deliver significant benefits for CAP strategic planning and monitoring. The integrated tools enhance policy coherence, evidence-based (data-driven) decision-making, and strategic foresight. Integrated toolboxes for evidence-based decisions are therefore crucial when MSs aim to ensure that different parts of the policy cycle, such as ex-ante analysis, monitoring and long-term visioning, fall into a single coherent framework. They can play a key role in complex tasks such as eco-scheme calibration, climate policy and nature conservation planning, and the alignment of agricultural policy with broader national strategies. This may be particularly important now that, following the proposal by the European Commission, traditional EU policies are being brought together within a single programming framework.

At the same time, the integrated toolbox approach suffers from limited accessibility, high requirements in terms of skills, time and data, and the need for careful adaptation to national contexts, which is both a strength and a weakness. While their added value in CAP strategic planning and foresight is clear, these limitations underscore the need for timely investment in data and modelling infrastructure, capacity-building within administrations and the research community, and sustained support from the European Commission and existing expert networks to share experiences and make such integrated approaches more widely usable across the EU.

A toolbox can be developed for various issues of strategic agricultural policy planning. It can focus on a more detailed evaluation and assessment of the effects of a single block of interventions or the achievement of a specific objective. It can also be used for a more comprehensive examination of individual phases of strategic planning, as demonstrated by vision building, ex-ante evaluation of policy reform, or assessing effects of changes in agricultural policy ex-post.

Integral toolbox support for strategic planning means that long-term trust has been established in the country among decision-makers, especially the administration, towards research teams and institutions who are able to support programming and monitoring with the development and implementation of tools. Knowledge and competencies are needed on both sides, in terms of both the developers and users of the toolboxes. The mutual recognition does not happen overnight, but can be built up over years. Therefore, a sufficiently high level of skills (and general education in policy affairs) and adherence to the principles of evidence-based policy making are also necessary.

Both CSs have shown that there exists a supply and demand for toolboxes related to agricultural policy programming in the EU, but they are not in widespread use across the EU, and experiences with their use vary greatly across the EU. The best indicator that there is a culture willing to make use of such toolboxes is the application of larger, more complex model tools for assessing the impact of agricultural policy options.

3. Conclusions and next steps

The document synthesises findings from 12 CSs that tested various tools to support CAP Strategic Plan (SP) design and monitoring. These tools fall into four main categories:

- *Quantitative tools* – mathematical simulation models (CAPRI, GLOBIOM) and experimental economics.
- *Qualitative tools* – participatory decision-making and stakeholder engagement methods.
- *Monitoring tools* – performance monitoring dashboards (PMT) and soil organic carbon mapping (SOC Map).
- *Integrated toolboxes* – combining multiple methods for Vision participatory building and eco-scheme impact assessment.

The assessment applied the *Tune-up Evaluation Framework* (TEF), assessing tools against five criteria: *accuracy, reliability, applicability, accessibility, and efficiency*. Insights were drawn from CS reports, factsheets, and internal workshops.

3.1. Key findings by tool category

Quantitative Tools

CAPRI and GLOBIOM offer robust, harmonised datasets and strong theoretical foundations. The tools are effective for EU-level policy analysis and multi-objective assessments (economic, environmental, climate). CSs have also demonstrated their usefulness for conducting impact assessments of various agricultural policy issues at the MS level, provided that the conditions are met. However, the application of such models at MS level is associated with various constraints: national adaptation requires time for development, specialised expertise, continuous use to develop experience and reduce marginal costs, and a stream of ongoing stable funding. Experimental economics provides behavioural insights, enabling more precise design of individual measures, but is rarely commissioned by policymakers. Academic expectations as to how these tools are used differ considerably from the reality of use in the process of designing measures.

If decision-makers decide to use large complex models, it is necessary to invest in the long term in cooperation with research teams (international and domestic) that manage these models. The best method is therefore to carry out continuous impact assessments under changing policy, market, and climate conditions. We also highly recommend the use of partial-equilibrium models (e.g., AGMEMOD) or farm-level simulations for MSs. These tools also enable impact assessments, but can be implemented primarily with national research infrastructure, since they are easier to adapt to national circumstances and priorities. Even with such an approach to support for policy making, it is necessary to invest in national research groups and to ensure cooperation between them and the ministry on an ongoing basis.

It should also be emphasised that it is necessary to formalise impact assessment requirements to promote evidence-based policymaking, including at the national level. The EU legislative framework can contribute to this, but national decisions on what is required can be even more effective. For comprehensive policy evaluation, it is necessary to use different models and link them together in a meaningful way (see the approach in integral case studies). We also recommend familiarising decision-makers with the possibilities of various experimental economics techniques. To start with, it is sufficient to make greater use of surveys and qualitative approaches.

Qualitative Tools

In the CSs, we tested several qualitative and participatory tools, including the IOI matrix for assessing the potential impact and coherence of interventions in relation to the set objectives, intervention logic, and a prioritisation voting tool. These tools support evidence-based CAP decision-making by structuring reasoning, eliciting stakeholder preferences, and clarifying intervention logic. They help policymakers and stakeholders understand and communicate trade-offs, coherence, and causal pathways, when designing or adjusting their CAP Strategic Plans. These tools are characterised by high flexibility (they can be easily adapted to needs), transparency, and low cost. They are useful for needs assessment, prioritisation, coherence analysis, and intervention design. Quick implementation is possible (workshops can be carried out over a small number of days). The tools are mostly low-cost: the main operational needs revolve around either collecting information through desk research and/or expert elicitation, or finding good facilitators.

Qualitative tools also have their limitations: they depend on stakeholder diversity and the quality of preparation and facilitation around the stakeholder engagement process. Results can reflect power imbalances in the stakeholder community or administrative biases in the decision-making body implementing the tools. The main limitations of qualitative tools (which can equally be viewed as their main strengths) relate to subjectivity, dependence on participant diversity, and the quality of documentation made available to participants and the broader policy community, which enhances the legitimacy and reliability of stakeholder engagement processes. However, they are certainly ideal for MSs with limited modelling capacity and tight budgets, or as a complement to empirical results of quantitative tools that enable structured reasoning. However, it cannot be overstated that the quality of their results strongly depends on the quality of the process.

For wider use, these tools also need to be well-equipped with instructions, operational standards need to be set, and moderators and participants need to be trained to ensure effective participation.

Monitoring Tools Examples

In CSs, we studied two new tools that can improve policy monitoring: PMT (Performance Monitoring Tool) and SOC Map (Earth observation and machine learning to estimate Soil organic carbon). These tools are designed to support the CAP's performance-based governance by integrating administrative, environmental, and Earth-Observation data from various sources to enable systematic tracking of policy results and environmental outcomes.

The advantage of such tools is that they enable systematic tracking of CAP results and environmental outcomes. However, such data are not easy to obtain, requiring high initial investment in IT infrastructure, data harmonisation, and technical expertise. Accessibility is lower, and administrations are mostly cautious when it comes to using such tools. Investments in this type of infrastructure are also necessary, often along with new legal frameworks that enable data acquisition or at least sharing. Monitoring tools can offer substantial long-term value for evidence-based CAP implementation, but their use is currently quite limited and left to their main proponents, i.e., research institutes and IT companies. The reasons for this are various constraints, ranging from technical capacity requirements, potentially high initial costs, data interoperability and privacy challenges, and dependence on specialised expertise and formal procedures for their integration.

Given the undoubted advantages they offer, MSs and EU-level decision-makers should be more courageous/ambitious in opting for new digital approaches when defining agricultural policy indicators. This, of course, requires investment in data infrastructure, institutional coordination, and continuous methodological development.

Integrated Toolboxes Case Studies

The Dutch CS focuses on eco-scheme impact assessment, while the Slovenian CS demonstrated the construction and use of a toolbox supporting vision building and strategic foresight. Integrated toolboxes combine several methods, quantitative models, monitoring systems, foresight, and participatory approaches into a coherent analytical framework. These are two different cases: the Dutch CS analysed the main features and potential impacts of the eco-scheme, while the Slovenian case study analysed the use of various tools and approaches for a more evidence-based preparation of the Vision as an integral part of the preparation of the new round (2028+) of CAP strategic planning. Both cases combine quantitative models, monitoring systems, and participatory approaches for balanced analysis and enhance policy coherence and strategic foresight. The use of such combined tools has high entry barriers. There are significant requirements in terms of skills, long setup (lead) time, and significant financial costs. Therefore, national research teams must be highly qualified for this type of work, and decision-makers must indicate a demand for such an approach. This requires strong institutional trust and long-term collaboration between administrations and research teams.

Both case studies show that integrated toolboxes demand long-term investment, expert teams, and institutional support, but deliver significant benefits for CAP strategic planning and monitoring. Integrated toolboxes can be developed for various issues of strategic agricultural policy planning. They can focus on a more detailed evaluation and assessment of the effects of a single block of interventions or the achievement of a specific objective. They can play a key role in challenging tasks such as calibrating complex policy mixes (here eco-schemes, but potentially also decoupled direct payments, climate policy, nature conservation planning, etc.), and the alignment of agricultural policy with broader national strategies. They can also be used for a more comprehensive examination of individual phases of strategic planning, as demonstrated by vision building, or for ex-ante evaluation of policy reform, and ex-post assessments of the effects of changes in agricultural policy.

3.2. Final assessment

Methodology and quality of case studies

The CSs applied a variety tools for end users. Both of the integrated toolboxes (Slovenia, Netherlands) have already been directly included in strategic planning for CAP post-2027. The tools in Germany, Hungary, France, Ireland and in both Spanish cases, were tested with national/regional administrations. In Poland, Greece, Italy, and the Czechia, the tools were tested with other stakeholders involved in strategic planning (NGOs, research institutions, mix of representatives).

From the perspective of direct use of tools, we can report on the following approaches in the CSs. In most CSs (The Netherlands, Ireland, Germany, Hungary, France, Czechia, Spain), already established tools were tested. In the second CS in Spain (use of new data for the CAPRI model) and Slovenia (new approach to vision building), the research teams attempted to develop or refine tools. Where there was no interest among end users for direct testing, a discussion and promotion of the tools was carried out (Greece, Italy, and Poland). Across the case studies, a consistent picture emerges regarding how policy analysis and monitoring tools interact with national administrative realities.

The Czech and Hungarian modelling CS show that advanced quantitative impact assessment tools can generate robust insights for CAP design, but only when they are embedded in stable institutional arrangements. In both cases, the lack of formalised national impact assessment frameworks and limited demand from ministries meant that modelling remained largely ad hoc and dependent on individual research capacity rather than becoming part of routine policymaking.

The Polish and Irish qualitative CS demonstrate that simpler tools can be more feasible and effective in practice, especially where time, data and analytical capacity are limited. These tools helped structure policy discussions, improve transparency and support shared understanding, but their effectiveness depended heavily on stakeholder participation and clear facilitation.

The French experimental CS illustrates the limits of technically demanding methods when institutional support and resources are insufficient. Despite methodological soundness, the tool's influence on decision-making remained limited due to low participation, time constraints and weak incentives for policymakers to engage.

The Italian and Greek monitoring pilots highlight both the potential and the challenges of building integrated monitoring systems. While the pilots demonstrated conceptual feasibility, they also revealed significant barriers related to data access, interoperability, and, more importantly, institutional coordination and administrative capacity. These cases underline that monitoring tools cannot succeed without institutional readiness and long-term investment in data governance.

The Spanish SOC mapping CS shows how innovative, EO-based methods can complement traditional monitoring approaches and reduce long-term data collection burdens, particularly at larger scales. At the same time, it confirms that such methods remain specialised and depend on high-quality ground data, technical expertise and clear policy demand.

Finally, the Dutch and Slovenian integral case studies demonstrate the benefits of integrated analytical ecosystems that combine multiple tools, data sources and participatory processes. These cases show that where analytical capacity is institutionalised and supported by long-term infrastructure and expertise, tools are more likely to be (re-)used, trusted and embedded in policy cycles.

In general, we might have expected greater interest in the tools among key decision-makers (civil servants employed within ministries of agriculture) in most MSs than was actually the case. This may be due to insufficiently strong links between the members of the project consortium and their respective ministries, which made fostering engagement around the tools challenging. Alternatively, it may actually be the case that there is a low actual interest in these tools among key end users. It is difficult to assert which explanation is correct. In any case, we did not observe particular enthusiasm for the tools among end users. Even research groups with long-standing and stable relationships with ministries had difficulty engaging the interest of the other side. This suggests that **end-user (civil servant) interest is the biggest challenge for the further use of tools and thus for strengthening the evidence-based and participatory approach to agricultural policy design and monitoring**.

The fact that the clients' interest in collaborating with research groups is very limited in most cases is a cause for concern. Only research groups with long-standing relationships with the ministry on strategic planning have

managed to establish a productive relationship with the end-user. The main reason for the project's only partially fulfilled ambitions is thus the gap between the project's expectations and the willingness and capacity of state administrations, as key users, to modify established processes of strategic planning and monitoring. Among other factors (see e.g. OECD, 2015; Suazo-Galdames et al., 2025; Hinrichs-Krapels et al., 2020), this can be attributed to the fact that, in many cases, project partners were operating from a supply-driven 'push' perspective, rather explicit policy demand, a well-documented problem in evaluations of research and CSA projects, and policy-application of research in general (Weiss, 1979; Cash et al., 2003, Szüdi et al., 2023).

However, despite certain difficulties in implementation and contacts with end-users, we believe that the CSs have achieved their fundamental goal: to test methods and alert administrations to their potential and the available means to improve strategic planning and monitoring.

Final conclusions and recommendations

What lessons can be drawn from conducting the case studies? We would highlight the following:

- Investment in capacity building and data and tools infrastructure is essential for wider adoption of advanced tools.
- Institutional trust and collaboration between policymakers and researchers, as well as stakeholders, are critical for success.
- Integrated approaches deliver the highest strategic value, but require sustained support from the EU and expert networks.
- Readiness for evidence-based policymaking varies across MSs; the use of tools, especially complex modelling tools and integral toolboxes, is still rare.

Despite certain gaps in the use of tools in most of the MSs studied, the CSs demonstrated the importance and necessity of using tools to improving strategic planning. Case studies also highlighted the utility of incorporating objective facts, data, and models for better planning. We can particularly highlight the use of participatory methods, which are simpler to apply and incorporate in decision-making process, and whose use should in fact be further expanded – it is at least implicitly demanded by the requirement of the CAP legal framework to engage with stakeholders.

It is therefore crucial to achieve a stronger understanding of the needs, benefits and possibilities of using tools by end user and support administrations by providing with additional knowledge and capacity necessary to conduct high-quality (evidence-based and legitimate) strategic planning. It is also necessary to raise competencies for dialogue on both sides of the science-policy interface and possibly even strengthen the formal framework for the use of tools at the EU level.

The knowledge and data requirements needed for the proper application of most tools to support evidence-based policymaking (except participatory tools) are very complex. It therefore is important to adapt the tools to national contexts, which requires specific knowledge and continuous investment in national research groups that can provide support, as well as thoughtful consideration of the possibility of improving data sources and their interoperability to prevent an unmanageable increase in administrative burden to both providers of data and data managers.

We conclude the report with the following recommendations, directed primarily at EU policymakers:

- Encourage formalised impact assessment frameworks at national level.
- Promote simpler, modular tools and use of participatory tools for countries with limited resources.
- Support training and knowledge exchange through EU networks.
- Invest in long-term infrastructure for monitoring and integrated analysis, while avoiding increasing administrative burden.
- Foster participatory processes to improve legitimacy and inclusiveness of CAP decisions.

These recommendations need to be carefully considered, upgraded, and developed in the final output of the project, Deliverable 5.3. - The Handbook.

4. References

Baehler, K. (2007). Intervention Logic/ Program Logic: Toward Good Practice. In J. Wanna (Ed.), Improving Implementation: Organisational Change and Project Management (1st edn). ANU Press. https://doi.org/10.22459/II.02.2007.14

Barreiro-Hurle, J., Fellmann, T., & M'barek, R. (2024). Modelling in Support of Better Agricultural and Food Policies: The JRC's Integrated Agro-economic Modelling platform (iMAP). *EuroChoices*, *23*(1), 43–53. https://doi.org/10.1111/1746-692X.12421

Brams, S. J., & Fishburn, P. C. (2002). Chapter 4 Voting procedures. In Handbook of Social Choice and Welfare (Vol. 1, pp. 173–236). Elsevier. https://doi.org/10.1016/S1574-0110(02)80008-X

Britz, W., Lacombe, D., & Moro, D. (2021). The FARMDYN model: A modular bio-economic farm model. *European Review of Agricultural Economics, 48*(4), 655–692. https://doi.org/10.1093/erae/jbab009

Cagliero, R., Licciardo, F., & Legnini, M. (2021). The Evaluation Framework in the New CAP 2023–2027: A Reflection in the Light of Lessons Learned from Rural Development. Sustainability, 13(10), 5528. https://doi.org/10.3390/su13105528

Cagliero, R., Mazzocchi, G., Monteleone, A., Pierangeli, F., Manzoni Di Chiosca, P., & Romano, E. (2022). A participative methodology for prioritising intervention logic in the design of the Italian CAP Strategic Plan. Rivista Di Economia Agraria, 77(3), 25–40. https://doi.org/10.36253/rea-13717

Cash, D. W., Clark, W. C., Alcock, F., Dickson, N. M., Eckley, N., Guston, D. H., Jäger, J., & Mitchell, R. B. (2003). Knowledge systems for sustainable development. Proceedings of the National Academy of Sciences, 100(14), 8086–8091. https://doi.org/10.1073/pnas.1231332100

Chantreuil, F., Hanrahan, K., & van Leeuwen, M. (Eds.). (2012). *The AGMEMOD model: Regional and EU agricultural market outlook.* JRC Scientific and Technical Reports. https://doi.org/10.2791/22068

Erjavec, E., Donnellan, T., van Leeuwen, M., & Salputra, G. (2018). CAP modelling for policy analysis: Lessons learned and challenges. *EuroChoices, 17*(3), 32–37. https://doi.org/10.1111/1746-692X.12217

European Commission. (2020). *A long-term vision for the EU's rural areas.* https://ec.europa.eu/info/strategy/priorities-2019-2024/stronger-europe-world/long-term-vision-eu-rural-areas_en

European Commission. (2021). Public administrations in the EU Member States – 2020 overview. Publications Office. https://doi.org/doi/10.2887/793815

European Commission. (2022). *Implementing the New Delivery Model of the CAP.* https://agriculture.ec.europa.eu

European Commission. (2025/560). Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL Establishing the Conditions for the Implementation of the Union Support to the Common Agriculture Policy for the Period from 2028 to 2034. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52025PC0560

European Commission. (2025/565). Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL establishing the European Fund for economic, social and territorial cohesion, agriculture and rural, fisheries and maritime, prosperity and security for the period 2028-2034 and amending Regulation (EU) 2023/955 and Regulation (EU, Euratom) 2024/2509. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52025PC0565

FAO. (2021). *SAFA – Sustainability Assessment of Food and Agriculture Systems: Guidelines and Indicators.* https://www.fao.org/nr/sustainability/safa

Gonzalez-Martinez, A, R., Jongeneel, R., Kros, H., Lesschen, J.P., de Vries, M., Reijs, J., Verhoog, G. (2021). Aligning agricultural production and environmental regulation: An integrated assessment of the Netherlands. Journal of Land Use Policy, Volume 105, June 2021, 105388

Hargreave, M., McNeill, A., & Jones, S. (2022). Challenges in integrating environmental indicators into agricultural monitoring. *Environmental Science & Policy, 136*, 171–182. https://doi.org/10.1016/j.envsci.2022.06.008

Hazell, P., & Norton, R. (1986). *Mathematical Programming for Economic Analysis in Agriculture.* Macmillan. https://doi.org/10.1007/978-1-349-18500-6

Hinrichs-Krapels, S., Bailey, J., Boulding, H., Duffy, B., Hesketh, R., Kinloch, E., Pollitt, A., Rawlings, S., Van Rij, A., Wilkinson, B., Pow, R., & Grant, J. (2020). Using Policy Labs as a process to bring evidence closer to public policymaking: A guide to one approach. Palgrave Communications, 6(1), 101. https://doi.org/10.1057/s41599-020-0453-0

Ji, X., Purushothaman, B., Prasad, R. V., & Aravind, P. V. (2024). Developing a digital mapping of soil organic carbon on a national scale using Sentinel-2 and hybrid models at varying spatial resolutions. Ecological Indicators, 167, 112654. https://doi.org/10.1016/j.ecolind.2024.112654

Kelly, R. A., Harou, J. J., López, A., et al. (2013). Selecting among five common modelling approaches for integrated environmental assessment. *Environmental Modelling & Software, 59*, 52–73. https://doi.org/10.1016/j.envsoft.2014.05.025

Linstone, H. A., & Turoff, M. (1975). The delphi method. Addison-Wesley Reading, MA.

Macleod, C. (2016). How can logic modelling improve the planning, monitoring and evaluation of policy measures and wider interventions for multiple benefits?

Marta-Pedroso, C., Laporta, L., Gama, I., & Domingos, T. (2018). Economic valuation and mapping of Ecosystem Services in the context of protected area management (Natural Park of Serra de São Mamede, Portugal). One Ecosystem, 3, e26722. https://doi.org/10.3897/oneeco.3.e26722

Matthews, A. (2015). The intervention logic of the CAP. Capreform.Eu. http://capreform.eu/the-intervention-logic-of-the-cap/

MEF4CAP Consortium. (2023). *Monitoring and Evaluation Framework for the CAP – Final Reports.* https://mef4cap.eu

Miles, I. (2010). The development of foresight: A retrospective view. Technological Forecasting and Social Change, 77(9), 1448–1455. https://doi.org/10.1016/j.techfore.2010.06.009

Miles, I. (2010). The development of foresight: A retrospective view. *Technological Forecasting and Social Change, 77*(9), 1448–1455. https://doi.org/10.1016/j.techfore.2010.06.009

Mortelmans, D., Fickel, T., Ott, E., Turkelboom, F., & Mehring, M. (2020). Policy Coherence analysis (PolCA): Methodological approach.

OECD. (2015). Scientific Advice for Policy Making: The Role and Responsibility of Expert Bodies and Individual Scientists (OECD Science, Technology and Industry Policy Papers No. 21; OECD Science, Technology and Industry Policy Papers, Vol. 21). https://doi.org/10.1787/5js33l1jcpwb-en

OECD. (2017). *Core Set of Environmental Indicators.* https://doi.org/10.1787/9789264278935-en

OECD. (2019). *Strategic Foresight for Better Policies.* OECD Publishing. https://doi.org/10.1787/9789264311831-en

OECD. (2020). Innovative citizen participation and new democratic institutions. OECD Publications Centre. https://www.oecd.org/content/dam/oecd/en/publications/reports/2020/06/innovative-citizen-participation-and-new-democratic-institutions_11aa2baf/339306da-en.pdf

Patton, M. Q. (2015). *Qualitative Research & Evaluation Methods* (4th ed.). Sage. https://us.sagepub.com/en-us/nam/qualitative-research-evaluation-methods/book232962

Petersen, E. H., Fraser, T., & Hone, P. (2018). Modelling agricultural systems for policy analysis. *Agricultural Systems, 162*, 1–3. https://doi.org/10.1016/j.agsy.2018.02.009

Potori, N., Himics, M., Witzke, P., Szabo, Z., Savoly, J., Egri, E. (2023): Socio-economic Implications of Banning Conventional Farrowing Crates in EU Pig Farming: A CAPRI-based Scenario Analysis, Studies in Agricultural Economics Volume 125 - Issue 3, pp 36-142, https://doi.org/10.7896/j.2584

Suazo-Galdames, I. C., Saracostti, M., & Chaple-Gil, A. M. (2025). Scientific evidence and public policy: A systematic review of barriers and enablers for evidence-informed decision-making. Frontiers in Communication, 10, 1632305. https://doi.org/10.3389/fcomm.2025.1632305

Szüdi, G., Bartar, P., Weiss, G., Pellegrini, G., Tulin, M., & Oomen, T. (2023). New trends in science communication fostering evidence-informed policymaking.

Tools4CAP Consortium (2023a). *Deliverable D1.1: Inventory of methods & tools used by MSs*. Horizon Europe Project Tools4CAP (Grant Agreement No. 101086311). https://www.tools4cap.eu/wp-content/uploads/2023/10/D1.1-Inventory-of-methods-and-tools.pdf

Tools4CAP Consortium (2023b). *Deliverable D1.2: Tools4CAP conceptual framework*. Horizon Europe Project Tools4CAP (Grant Agreement No. 101086311). https://www.tools4cap.eu/wp-content/uploads/2023/10/D1.2-Conceptual-framework.pdf

Tools4CAP Consortium (2024a). *Deliverable D1.3: Evauation of methods and tools for the design and monitoring of the CAP strategic plans.* Horizon Europe Project Tools4CAP (Grant Agreement No. 101086311). https://www.tools4cap.eu/wp-content/uploads/2024/03/Tools4CAP_Evaluation-and-benchmarking-of-methods-and-tools-D1.3.pdf

Tools4CAP Consortium (2024b). *Deliverable D1.4: Inventory of methods & tools used by MSs – Extended version*. Horizon Europe Project Tools4CAP (Grant Agreement No. 101086311). https://www.tools4cap.eu/wp-content/uploads/2025/03/D1.4-Inventory-of-methods-tools-used-by-MSs-II.pdf

UNDP. (2018). Foresight Manual: Empowered Futures for the 2030 Agenda. https://www.undp.org/publications/foresight-manual

UNDP. (2018). *Foresight Manual: Empowered Futures for the 2030 Agenda.* https://www.undp.org/publications/foresight-manual

van Ittersum, M. K., Ewert, F., Heckelei, T., et al. (2008). Integrated assessment of agricultural systems: A component-based framework. *Agricultural Systems, 96*(1–3), 150–165. https://doi.org/10.1016/j.agsy.2007.07.009

van Notten, P. (2005). Scenario development in times of discontinuity. OECD International Futures Project. https://doi.org/10.1787/9789264017726-en

Weiss, C. H. (1979). The Many Meanings of Research Utilization.

Yuzugullu, O., Fajraoui, N., Don, A., & Liebisch, F. (2024). Satellite-based soil organic carbon mapping on European soils using available datasets and support sampling. Science of Remote Sensing, 9, 100118. https://doi.org/10.1016/j.srs.2024.100118

Zhang, G., Liu, F., & Song, X. (2017). Recent progress and future prospect of digital soil mapping: A review. Journal of Integrative Agriculture, 16(12), 2871–2885. https://doi.org/10.1016/S2095-3119(17)61762-3

5. Appendices – Case Studies Factsheets

TOOLS4CAP

CASE STUDY FACTSHEET

GLOBIOM-CZ for Ex-Ante Analysis and Environmental Assessment

CZECHIA

Description and relevance

The Czech case study applied the nationalised GLOBIOM-CZ model to analyse **how CAP interventions interact with climate change, competitiveness, and environmental objectives**. Czechia is unique among EU Member States for allocating a particularly **high share of its CAP budget to environmental and climate priorities**, and the case study aimed to examine whether these investments deliver **balanced outcomes across production, emissions, land use, and resilience**.

The rationale was to provide evidence that helps **reconcile competing concerns**: farmers' need for economic viability and predictable conditions; policymakers' commitment to Green Deal targets; and society's expectations for climate mitigation and biodiversity protection. GLOBIOM-CZ was therefore used to **explore alternative "policy pathways,"** combining CAP instruments (direct payments and GAECs) with climate shocks to 2030 and 2050.

The case sought to **identify synergies, trade-offs, and risks**, providing Czech stakeholders—including ministries, research institutes, farmers, and regional actors—with a transparent analytical basis for preparing future CAP cycles and national climate-adaptation strategies.

ABOUT THE CASE STUDY

TOOL CATEGORY: Quantitative models

GEOGRAPHIC LEVEL: National

TYPE: Delivered ready for use + Showcase

Ing. Juliana Arbelaez Gaviria
prof. Mgr. Ing. Miroslav Trnka, Ph.D., *CS coordinator*
Petr Havlík, Ph.D.
Amanda Palazzo

VISIT THE CASE STUDY SITE!

Tools applied

- GLOBIOM-CZ (partial-equilibrium model with national parameters).
- Climate impact layers from EPIC-IIASA and CMIP6 (yield + water availability).
- Pillar I payments integrated as endogenous incentives.
- GAEC rules encoded as binding constraints (GAEC 1, 4, 7, 8).
- Scenario ladder (No CAP/DP-only/GAEC-only/Full CAP).
- Indicator Dashboard (IIASA Accelerator – forthcoming).

Funded by the European Union

Tools4CAP has received funding from the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement No. 101086311. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Results and practical innovations

The modelling demonstrated **strong interactions between CAP instruments and climate change**. GAEC 1 (permanent grassland) emerged as a **robust stabiliser of environmental performance**, consistently maintaining grassland shares above the national threshold even under strong warming. **Direct payments** alone increased emissions in some crops (e.g., potatoes, wheat), while the **full CAP package mitigated these effects through conditionalities**. Climate change **amplified spatial and commodity-specific vulnerabilities**, especially for maize, while creating relative advantages for crops like potatoes and rapeseed.

Economic indicators showed that production, trade, and competitiveness shift substantially **depending on policy mixes**. The case introduced **several innovations**: direct payments and GAECs were implemented inside the optimisation itself (not post-hoc); climate impacts were integrated globally for full EU comparability; and a transparent scenario ladder was designed to support communication with non-technical stakeholders. The approach created a **traceable workflow linking real policy levers to environmental and economic outcomes**, suitable for iterative policy refinement.

Benefits, limitations, and risks

The main benefit of GLOBIOM-CZ is its **ability to quantify policy trade-offs** in a transparent, reproducible manner that **combines climate science with economic modelling**. It provides Czech stakeholders with comparable results across scenarios, warming levels, and policy mixes. This helps identify when conditionalities are needed to **temper negative externalities of payments**, and where competitiveness opportunities lie.

However, limitations remain: GLOBIOM operates at sector level and **cannot capture farm-level heterogeneity or behavioural dynamics**; it also cannot fully represent **detailed eco-scheme rules** without complementing models.

Risks include **misinterpretation if results are detached from assumptions, incomplete stakeholder engagement** (particularly limited ministry participation), and **overreliance on model outputs** without validating them against national data. The study mitigated these risks through traceability, version control, and explicit documentation of uncertainties.

Still, successful application requires a **maintained baseline**, regular **expert validation**, and **clear communication** to avoid misuse of results.

Recommendations

Czechia should **institutionalise an ongoing partnership** between CzechGlobe (as national steward) and IIASA (as methodological steward) to maintain a continuously updated GLOBIOM-CZ baseline and enable rapid scenario testing when policy windows open. **Validation with analytical partners** such as ÚZEI should become routine, and future iterations should enforce country- and measure-level Pillar I budget caps, refine regional differentiation, and develop simple "scenario explorer" dashboards for practitioners and policymakers.

For deeper insight into distributional effects, **GLOBIOM-CZ should be linked with farm-level micro-simulation models**. Stakeholder engagement should be broadened—particularly with ministries—to secure early buy-in and ensure that results inform real CAP adjustments.

Finally, the Czech approach should be **replicated across other Member States** using the EU-wide GLOBIOM-EU framework, **enabling cross-country comparison of CAP mixes** under climate change and reinforcing evidence-based governance at both national and EU levels.

TOOLS4CAP

CASE STUDY FACTSHEET

IOI Matrix for Policy Coherence in Environmental Measures (Hesse)

GERMANY

Description and relevance

The German case study explored the use of the IOI (Intervention-Objective-Impact) matrix as a structured tool to **analyse the coherence of policy measures** in the federal state of Hesse. The objective was to identify how eco-schemes under Pillar I and agri-environment-climate measures (AECMs) under Pillar II **interact, overlap, or complement** one another in achieving **climate, water, and biodiversity goals**.

The case responds to a **recurring policy challenge in Germany's multi-level governance system**—ensuring consistency between national and regional programmes. By mapping instruments against objectives and expected impacts, the IOI matrix provides a transparent framework for **assessing synergies and trade-offs**.

The exercise aimed to demonstrate its applicability for **ex-ante design, mid-term review, and ex-post evaluation** of CAP interventions, while enhancing collaboration between ministries, paying agencies, and researchers.

ABOUT THE CASE STUDY

TOOL CATEGORY: Qualitative and participatory methods

GEOGRAPHIC LEVEL: Regional

TYPE: Retroactive, limited adoption for the last funding period for practical use for end users

ifls

Carla Wember, *CS coordinator*
Heike Nitsch
Simone Sterly

Results and practical innovations

The populated IOI matrix allowed researchers to **visualise the relationships between instruments and objectives**, revealing clusters of complementary actions and areas of potential duplication. For example, eco-schemes addressing nutrient management and water quality showed strong alignment with AECM measures but partial overlap in biodiversity objectives.

The structured approach helped to prioritise instruments with high environmental effectiveness relative to cost. The tool facilitated a **common language** for policymakers and analysts to discuss **evidence and identify policy gaps**.

The process itself became a **capacity-building exercise**, encouraging ministries to adopt more systematic planning practices. The innovation lies not only in the analytical structure but also in its ability to function as a **coordination tool bridging science and administration** within a complex governance framework.

Funded by the European Union

Tools4CAP has received funding from the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement No. 101086311. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Tools applied

- **IOI Matrix framework** for policy coherence analysis.
- **Desk research** and legal **mapping** of instruments.
- Expert **consultations** and validation **meetings**.
- **Qualitative scoring** of effectiveness and coherence.

Benefits, limitations, and risks

Key benefits include **enhanced transparency, reduced fragmentation, and better documentation** of policy logic. The IOI matrix supports evidence-based prioritisation and helps **visualise how different instruments jointly contribute to CAP's nine specific objectives**. It enables a structured dialogue between administrative levels.

However, it remains data-intensive and requires consistent **expert judgment** to fill qualitative scores, which introduces **subjectivity**. Limitations also stem from **time constraints** and **uneven stakeholder participation**.

Risks include **over-reliance on the matrix as a static tool** rather than a living process, and **potential fatigue** among officials if not clearly linked to decision-making. Nonetheless, the German experience proved that even a pilot can strengthen cross-departmental learning and accountability.

Recommendations

To mainstream the IOI matrix, Germany should **integrate it into regular planning and evaluation cycles** of CAP Strategic Plans. **Training officials** in qualitative assessment and **involving regional experts** can enhance accuracy and ownership.

The tool should evolve toward a **dynamic database connecting indicators, measures, and impacts**, reducing manual scoring. Future work could link it with **performance monitoring frameworks** (PMEF) and spatial uptake data to g**et a better grasp of effectiveness in practice**.

Finally, periodic workshops between science and policy teams should institutionalise the matrix as a **collaborative learning mechanism**, reinforcing policy coherence and efficiency in CAP implementation.

TOOLS4CAP

CASE STUDY FACTSHEET

SOCMap: Remote Sensing and Machine Learning for Soil Organic Carbon

SPAIN

Description and relevance

The Spanish SOCMap case study developed a methodology to estimate Soil Organic Carbon (SOC) using georeferenced soil samples, Earth Observation (EO) data, and Machine Learning algorithms. It aimed to provide Managing Authorities with a **practical, cost-effective method to evaluate interventions promoting carbon sequestration** within the CAP Strategic Plan.

Traditional soil sampling campaigns are accurate but **expensive and time demanding**; SOCMap addressed this by combining EO data from Sentinel-2 satellites with **ancillary datasets** such as topography, climate, and land use.

Implemented by ITACyL, the case produced a **prototype SOC map for Spain**, demonstrating how remote sensing and modelling techniques can complement field data and support CAP monitoring. The initiative showcased the value of open data and cloud computing for carbon accounting, aligned with EU goals on climate change mitigation and sustainable soil management, and offering a replicable model for other Member States.

ABOUT THE CASE STUDY

TOOL CATEGORY: **Data and monitoring systems**

GEOGRAPHIC LEVEL: **National**

TYPE: **Showcase**

INSTITUTO TECNOLÓGICO AGRARIO

Alberto Gutiérrez García, *CS coordinator*
Arturo Jiménez Beneite
David A. Nafría García

Tools applied

The case study applied several complementary tools and datasets to estimate SOC across Spain:

- A **georeferenced soil sample database**, integrating data from various regional and national institutions and farmers' organisations, provided the foundation for model calibration and validation.
- These samples were combined with **Sentinel-2 imagery from the Copernicus Data Space Ecosystem** (CDSE), allowing retrieval and processing of bare-soil reflectance composites. Additional auxiliary data, including topography, climate, and land-use layers, were incorporated to strengthen prediction accuracy.
- SOC estimations were produced through **Machine Learning models** (Random Forest, Cubist, and Partial Least Squares Regression) trained and validated using cross-validation.
- The workflow relied on **open-source GIS and cloud platforms**, such as QGIS and the OpenEO API, ensuring reproducibility, scalability, and full transparency.

Together, these tools formed an affordable methodology that can be easily adapted by public administrations for CAP-related monitoring.

Funded by the European Union

Tools4CAP has received funding from the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement No. 101086311. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Results and practical innovations

The pilot successfully generated national-scale SOC maps with quantified uncertainty. Combining **spectral, topographic, climatic and land-use predictors improved accuracy** compared to models based on limited sampling. Testing multiple algorithms enabled **identification of optimal configurations**, while new bare-soil reflectance composites enhanced consistency across time and space.

The process demonstrated that open data and software could produce **reliable outputs with limited licensing costs and dependence on proprietary solutions**. Practical innovations included modular workflows, systematic cross-validation, and the incorporation of uncertainty maps to communicate model reliability.

These advances strengthened confidence in remote sensing-based soil assessments. The methodology proved both feasible and operationally sound, demonstrating that digital, open-source approaches can support evidence-based CAP evaluation and inspire methodological harmonisation across the EU.

Benefits, limitations, and risks

Key benefits of SOCMap were **scalability, affordability, and replicability**. By using open data and, to the possible extend free software, the methodology reduced monitoring costs and supported data-driven evaluation of carbon farming practices. Once models were trained, outputs could be updated quickly, offering a flexible tool for continuous CAP monitoring. **Open-source tools** also made the system accessible for public institutions, promoting transparency and collaboration.

However, limitations included **uneven sample coverage, noise** from residual vegetation, and the **limited interpretability** of complex models.

Risks involved **misuse of outputs** without proper understanding of uncertainty or applying models outside their training domain.

The study also noted the importance of multidisciplinary expertise in GIS, EO, and soil science to ensure sound implementation. Despite these challenges, SOCMap proved invaluable for establishing **baselines**, **guiding soil-management strategies** and **strengthening evidence-based policymaking** within CAP.

Recommendations

The European Union should **invest in harmonising soil sample datasets**, expanding field sampling, and establishing technical standards for SOC mapping. Common data formats, quality protocols, and interoperable geodatabases would enhance comparability and ensure reliable monitoring. **Training** staff in GIS remote sensing and data management will promote institutional uptake and long-term sustainability. Integrating SOCMap results into CAP monitoring frameworks and linking them with EU initiatives such as the Soil Observatory would enhance **consistency and transparency**.

Further research should **explore radar and thermal sensors to reduce cloud interference**, as well as hybrid models combining physical and statistical approaches to enhance accuracy. SOCMap demonstrates how digital innovation and open data can make carbon monitoring both feasible and policy-relevant, fostering harmonised approaches across Member States.

TOOLS4CAP

CASE STUDY FACTSHEET

CAPRI & Digital Farm Book Data for Pesticide Use Evaluation

SPAIN

Description and relevance

This Spanish case study explored integrating Farm Management Information System (FMIS) data—representing voluntary Digital Farm Book entries—into the CAPRI model to improve policy analysis of pesticide reduction strategies. The aim was to **test the technical feasibility of using real farm-level data** to replace or refine traditional priors in agro-economic models, thereby improving the **accuracy and policy relevance of ex-ante assessments**.

Conducted by ITACyL and EuroCare, the study examined how data collected by a Farm Manangement Information System (FMIS) on pesticide applications from several Spanish farms could feed into CAPRI's structure to **evaluate production costs, environmental impacts, and compliance** with EU pesticide reduction targets.

The work addressed a core policy question: **how digitalisation and data interoperability can enhance CAP monitoring and decision support**.

ABOUT THE CASE STUDY

TOOL CATEGORY: **Integral tool (multi-tool approaches)**

GEOGRAPHIC LEVEL: **Regional**

TYPE: **Trial run**

INSTITUTO TECNOLÓGICO AGRARIO

EuroCARE AGRICULTURAL POLICY RESEARCH

Alberto Gutiérrez García, *CS coordinator*
David Schaefer
David A. Nafría García
Peter Witzke

Results and practical innovations

The integration of parcel-level records into CAPRI demonstrated **its flexibility** and the **analytical value of farm-generated data**. The main strength lies in the granularity of parcel-level information, which allows models to capture crop- and region-specific patterns that are otherwise invisible in aggregate statistics. Although focused on pesticides, the lessons extend to fertilisation, irrigation and livestock treatments. Additionally, the workflow successfully **anonymised, cleaned, and scaled FMIS data to national totals**, creating robust priors for CAPRI.

Innovations included **correction of implausible values**, introduction of **Highly Hazardous Pesticide indicators**, and development of **scalable data pipelines**. The exercise highlighted the **potential of Digital Farm Books as a new evidence source** for CAP assessment once reporting becomes mandatory.

Funded by the European Union

Tools4CAP has received funding from the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement No. 101086311. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Tools applied

- **Farm Management Information System** (FMIS) dataset from Digital Farm Books.
- **CAPRI** agro-economic partial-equilibrium model.
- **Scenario simulations** for pesticide use and hazardous substance bans.
- Harmonisation with **EUROSTAT sales data** and **JKI priors**.

Benefits, limitations, and risks

The main benefit was the enhanced realism of the model outputs, helping to **bridge the gap between actual farm practices and macro-level policy analysis**. The case study demonstrated how FMIS data can **strengthen national representation within EU-scale models**.

Policymakers gain value from linking FMIS data to agro-economic models, as this provides a more reliable foundation for designing and monitoring CAP measures. The case study offers a **unique dataset for testing scenarios** with greater realism and for examining structural differences across crops and regions. Its insights support benchmarking efforts and help **identify opportunities for innovation**.

However, limitations included **incomplete coverage** (with a bias toward cereals), **voluntary data reporting**, and various **data-quality challenges**. Additional risks relate to **privacy management** and potential **biases** if datasets remain unrepresentative.

Despite these constraints, the case study demonstrated a **replicable workflow** that connects digital agriculture with policy modelling and offers **lessons for future integration** of similar data sources across the EU.

In summary, incorporating FMIS data into modelling tools provides a cost-effective approach to aligning agricultural policy with environmental objectives and farm-level viability.

Recommendations

The main recommendation from this case study is that **Digital Farm Book reporting should be formalised** by developing anonymisation standards, and establishisng cooperation between data providers, administrations, and modellers. The key element is building **trust with farmers through transparency and simplified reporting** which, in turn, will ensure data quality.

The **CAPRI-FMIS framework** should be extended to **include fertiliser and irrigation data** to support integrated sustainability analysis.

Investments in capacity-building and IT infrastructure will enable other Member States to replicate this digital-model interface, transforming evidence-based agricultural policymaking under the CAP.

TOOLS4CAP

CASE STUDY FACTSHEET

Experimental Evaluation of Agri-Environment-Climate Measures (AECM)

FRANCE

Description and relevance

The French case study focused on testing innovative evaluation approaches for agri-environment-climate measures (AECMs) in the Occitanie region. This measure required farmers to implement diversified protein crop rotations and reduce external feed inputs, with the dual objective of **improving environmental performance and strengthening regional feed autonomy**. It addressed the persistent challenge of measuring the **behavioural and economic factors driving farmers' participation in schemes** designed to enhance environmental performance.

The goal was to assess **how farmers perceive, understand, and respond to policy attributes**, while exploring how behavioural insights can improve the design and acceptance of AECMs. This work is particularly relevant to the current CAP framework, which emphasises results-based and evidence-driven policies. By introducing experimental and behavioural evaluation tools, the case contributes to **enhancing the analytical capacity** of Managing Authorities and to **better aligning policy instruments with farmers' motivations** and local production realities.

ABOUT THE CASE STUDY

TOOL CATEGORY: **Quantitative models**

GEOGRAPHIC LEVEL: **Regional Occitanie**

TYPE: **Full adoption**

ORÉADE-BRÈCHE
Ingénierie de l'Environnement et du Développement

Pauline Mace, *CS coordinator*

Tools applied

The tool initially selected for the case study was the **Discrete Choice Experiment (DCE)**, intended to analyse how specific **attributes of agri-environmental schemes influence farmers' decision** to participate. However, due to implementation constraints encountered during the project, the approach was revised and the DCE was **replaced by the Theory of Planned Behaviour (TPB)**. This second tool focuses on external factors shaping farmers' intentions, such as **social norms, perceived behavioural control and personal attitudes**.

Although grounded in different theoretical frameworks, both tools share the objective of identifying the determinants of farmers' decision-making processes. Their combined use ultimately enabled the design of a **survey**, sent to **273 farmers**, that integrated elements from both approaches (retaining some attributes originally defined in the DCE), thereby enhancing the relevance and explanatory power of the data collected compared with a conventional survey.

Funded by the European Union

Tools4CAP has received funding from the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement No. 101086311. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Results and practical innovations

Originally conceived as a DCE to **estimate farmers' willingness to accept scheme conditions and payment levels**, the case study evolved toward a **behavioural framework** under data and legal constraints. Using the **Theory of Planned Behaviour**, researchers surveyed a representative sample of farms to capture **attitudes, perceived control, and social norms** affecting enrolment decisions.

Results highlighted the **decisive role of building on practices that align with existing systems or show proven effectiveness in scheme design, alongside farmers' trust in institutions and perceived administrative simplicity**. Preliminary statistical associations indicated that perceived administrative simplicity and existing familiarity with protein crops were among the strongest predictors of positive enrolment intentions. The work also revealed both the potential and limitations of combining behavioural insights with economic evaluation under real-world constraints: while the integration of TPB elements offered a more realistic understanding of farmers' motivations, the inability to implement a full DCE and the reliance on qualitative data due to a small sample reduced analytical depth.

Nonetheless, the pilot illustrated that adapted evaluation methods can still yield meaningful insights, even without monetary attributes, provided that problem framing is clear and stakeholder involvement is consistent. Overall, the experience suggests that **simpler, more participatory approaches may be better suited when time, data, or resources are limited**, while still informing the refinement and design of future schemes.

- Discrete Choice Experiment (DCE)
- Theory of Planned Behaviour (TPB) framework
- Stakeholder survey (273 farm participants)
- Behavioural analysis and preference modelling

Benefits, limitations, and risks

The French case demonstrated the practicality and policy value of behavioural evaluation methods. Its major benefit lies in **broadening the analytical scope of CAP monitoring** beyond pure economic rationality to include **perceptions, social dynamics, and values**. When implemented under suitable conditions, these methods offer actionable inputs for designing or adjusting schemes and anticipating participation drivers.

However, the experience also revealed important limitations: the absence of a monetary attribute, a **small and seasonally constrained sample, reduced analytical depth**, and **comparability**. The demanding design, **need for early stakeholder engagement**, and **significant time and expertise** led to further **constrained feasibility**.

Risks involve **misinterpretation of behavioural metrics** if not coupled with economic indicators, and that the effort invested may outweigh the direct policy value in resource-limited contexts.

Still, the case proved that behavioural approaches can **complement standard evaluations** and **increase the relevance of results** to policymakers. Their robustness, however, depends on realistic planning and sufficient resources, and suggests that simpler participatory methods may be more appropriate when these conditions are not met.

Recommendations

Future evaluations should **combine economic and behavioural approaches**, merging DCEs and TPB models into hybrid frameworks that can **quantify both financial and attitudinal factors**. Scaling up such methods requires close collaboration between regional agencies, research institutions, and farmer organisations to **secure data and stakeholder trust**. Establishing early data-sharing agreements between administrations and research teams would prevent methodological downgrades during implementation. Simplifying survey design and adopting digital platforms could expand participation and timeliness.

The findings also call for **capacity-building within Managing Authorities** to interpret behavioural and economic data. Ultimately, integrating behavioural evaluation into mainstream CAP monitoring could **enhance both the quality of evidence and the social legitimacy of environmental measures**.

TOOLS4CAP

CASE STUDY FACTSHEET

Policy Monitoring Tool for CAP Performance and Evaluation

GREECE

Description and relevance

The Greek case study tested the **Policy Monitoring Tool** (PMT) as an integrative platform to improve performance evaluation and data-driven decision-making for CAP Strategic Plan implementation. The PMT was designed to connect digital agriculture systems at parcel level, Earth Observation data, and administrative records into a single analytical framework. The objective was to **assess the feasibility of combining existing monitoring mechanisms with new digital data flows** to improve the timeliness and transparency of CAP performance reporting.

Greece, like many Member States, faces **structural challenges**: fragmented data systems, limited interoperability, and low stakeholder engagement. This case was conceived as a proof-of-concept to **show how modern data integration could enhance both national and EU-level performance reviews**. It focused on linking land-based indicators, eco-scheme participation, and environmental outcomes in a user-friendly and replicable format.

ABOUT THE CASE STUDY

TOOL CATEGORY: **Data and monitoring systems**

GEOGRAPHIC LEVEL: **Regional**

TYPE: **Trial run + showcase**

NEUROPUBLIC

Georgios Charvalis, ***CS coordinator***
Odysseas Sekkas

Tools applied

The **Policy Monitoring Tool** tested provides policymakers with a visual policy monitoring dashboard, allowing the generation of analytical **reports based on queries**, including summary tables and graphical charts. The tool integrates and renders data at a regional level (Local Administrative Units – LAUs) derived from **anonymised and aggregated in-situ pilot test cases**. This includes information from parcels utilizing Digital Agriculture Technological Solutions (DATSs), parcels not employing DATs, farm calendars, digital logs etc.

Earth Observation (EO) data products such as crop classification maps and European land use data (e.g. **ESA WorldCover)** are used as inputs, with open EU datasets (e.g. EEA, FAO). These sources support the regional extrapolation of aggregated key indicators, variables, or thresholds that enable both qualitative and quantitative comparisons of regional policy performance.

From a wider perspective, the tool functionalities revolve around the following three pillars:

- **Generalised indicators tracking** (focusing on variables such as agrochemicals use, irrigation practices, costs etc.).

Funded by the European Union

Tools4CAP has received funding from the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement No. 101086311. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

- **DATS vs. non-DATS parcel performance** (evaluating the effectiveness of Digital Agriculture Technologies (DATSs) compared to traditional farming methods).
- **Regional benchmarking** (Integrating regional benchmark values/thresholds derived from various heterogeneous sources).

Results and practical innovations

The case successfully **demonstrated the ability of the PMT to integrate data streams from multiple sources and visualise indicator trends** relevant to the CAP's environmental and socio-economic goals. Through collaboration with national stakeholders, test datasets were processed to **estimate indicators such as fertiliser application, irrigation intensity and yield outputs per growing season per Local Administrative Unit region**.

The tool revealed how digital systems could **automate reporting processes and support ex-ante and ex-post evaluation**. It also demonstrated that indicator-based monitoring can **complement field inspections**, lowering administrative burdens.

Practical innovations included **combining Earth Observation with in-situ data**, developing simple visual analytics, and establishing an early workflow for institutional cooperation. Although full implementation was limited by time and data access, the project **validated PMT's scalability for future CAP cycles**.

Benefits, limitations, and risks

The primary benefit was **the operational demonstration of an integrated monitoring solution** that can enhance evidence-based governance. It helps administrations **move from compliance-based to performance-based reporting**, aligning with the New Delivery Model of the CAP. Benefits include **cost efficiency, improved transparency**, and **higher data consistency**.

However, the pilot faced **significant limitations** due to **incomplete data availability, weak interoperability** between agencies, and the **lack of a formal data-sharing framework**. Risks include **system redundancy if not institutionalised**, and **data privacy issues** under evolving EU regulations.

An **important finding** is that data scalability uncertainties are greater when lack of complementary datasets exist, while the lack of interoperability with official databases (e.g. IACS) hindered policy-level validation.

Despite these constraints, the PMT provided a **replicable template adaptable** to various Member States.

Recommendations

To maximise its potential, Greece should **formalise institutional cooperation** around digital data management, establish legal protocols for data exchange, and train staff in implementing and using analytical tools.

Future iterations should **link PMT indicators directly to the CAP performance review framework** and integrate **farm-level feedback** mechanisms. **Investment in national databases and capacity-building** is essential to sustain long-term benefits. As data systems mature, the PMT can become a cornerstone for real-time monitoring, supporting both national evaluation and EU-level reporting obligations.

TOOLS4CAP

CASE STUDY FACTSHEET

CAPRI Modelling for Ex-Ante Policy Analysis

HUNGARY

Description and relevance

The Hungarian case study applied the CAPRI (Common Agricultural Policy Regionalised Impact) model to **analyse the economic and environmental implications of increasing coupled payments** for protein crops, particularly soybeans. The case aimed to demonstrate how quantitative models can **inform evidence-based decision-making** during the design of the CAP Strategic Plan.

This exercise responded to Hungary's need to **evaluate trade-offs between income support**, **production diversification**, **self-sufficiency** and **environmental sustainability** in the protein crop sector. The study also aligned with broader EU objectives on protein self-sufficiency.

It tested how the CAPRI model, as a harmonised partial-equilibrium framework, can **simulate changes in farm production, trade, and emissions** at regional and EU scales. The ultimate objective was to demonstrate that **ex-ante modelling is capable to bridge the gap between national policymaking and EU-wide targets**, enabling more informed and transparent strategy formulation.

ABOUT THE CASE STUDY

TOOL CATEGORY: **Quantitative models**

GEOGRAPHIC LEVEL: **National**

TYPE: **Trial run**

Institute of Agricultural Economics

Norbert Potori, *CS coordinator*
János Sávoly
Zsolt Szabó

Results and practical innovations

Scenario simulations revealed that either an increase in coupled support without upper hectare limit or an eco-scheme-like program targeting protein crops could significantly **boost soybean cultivation** and **enhance domestic protein self-sufficiency** in Hungary. Production shifts were accompanied by moderate decreases in cereal areas, especially maize and wheat, but **overall farm income remained broadly stable**.

Environmental results showed that greenhouse gas emission benefits offer some indication of how targeted policy measures can contribute to sustainability objectives. The exercise validated CAPRI's capacity to **capture regional heterogeneity and feedback effects** between crop and livestock sectors. Increasing the supply of sustainable, locally produced soybeans to reduce dependence on imported soy would have positive repercussions to global deforestation, GHG emissions, and biodiversity through crop diversification.

Innovations included **refining national parameters** for Hungarian conditions and **integrating local policy assumptions** into CAPRI's EU-wide framework. The case strengthened Hungary's analytical capacity to use modelling for ex-ante policy assessment and offered insights on designing **future eco-schemes linked to protein crops**.

Funded by the European Union

Tools4CAP has received funding from the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement No. 101086311. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Tools applied

- CAPRI agro-economic model (partial-equilibrium framework).
- National datasets on existing support schemes and expert parameters.

Benefits, limitations, and risks

The key benefit was the **ability to quantify trade-offs** between economic viability and environmental sustainability using consistent, EU-comparable data. CAPRI enabled evidence-based discussions on targeted support while revealing **structural sensitivities in the crop mix**.

Limitations included **data harmonisation challenges, technical complexity**, the need for **trained modellers** and a certain level of technical capacity that sets a high entry price for Member States or institutions intending to utilize it for policy design. In addition, the model's aggregate representation of farms could not fully capture **micro-level adaptation** or behavioural responses, because the applicability of partial economic models, such as CAPRI, may diminish in policy contexts that demand a more holistic understanding of economic and environmental interactions.

Risks concerned **overreliance on model results** without accounting for policy uncertainty or complementary instruments.

Nonetheless, the exercise **improved institutional cooperation** between researchers and ministries and demonstrated how modelling tools enhance transparency in CAP planning.

Recommendations

The Hungarian Case Study revealed that Member States' investment in agricultural modelling capacity and setting up collaboration with CAPRI network partners is a viable strategy to address outstanding agro-societal priorities in the national CAP strategic planning.

For agricultural policy officers and ministry advisors, it offers clear insights into the impacts of policy choices, financial allocations, and sectoral challenges in meeting their mandates. Thus, **scenario analysis** should become a **routine step in policy formulation**, evaluating interventions across in income, competitiveness, and environmental metrics.

Future research should **integrate CAPRI with bio-economic or farm-level models to improve behavioural realism. Expanding** access to high-resolution data, ensuring **documentation of assumptions**, and **involving stakeholders** in interpreting results will make modelling outputs more robust and trusted for future CAP cycles.

TOOLS4CAP

CASE STUDY FACTSHEET

Intervention Logic for CSP Monitoring and Evaluation

IRELAND

Description and relevance

The Irish case study focused on developing an Intervention Logic (IL) of direct payment interventions to strengthen the monitoring and evaluation of the CAP Strategic Plan 2023–2027. The aim was to **link direct payment interventions with measurable socio-economic and environmental outcomes with regard to specific CAP objectives**, providing a more coherent structure for performance analysis.

The case was motivated by the need to complement the EU's Performance Monitoring and Evaluation Framework (PMEF) with **additional, nationally relevant indicators that reflect local farming realities**. The IL was designed as a visual and logical mapping tool to demonstrate how interventions contribute to CAP objectives, enhancing transparency and accountability in policy delivery.

ABOUT THE CASE STUDY

TOOL CATEGORY: Qualitative and participatory methods

GEOGRAPHIC LEVEL: National

TYPE: Real-life adoption

teagasc
Agriculture and Food Development Authority

Paula Palma Molina
Emma Dillon
Trevor Donnellan

Tools applied

- **Intervention Logic (IL) modelling framework**

The IL diagram was developed by reviewing scientific and grey literature, including Ireland's CSP 2023-2027 and SWOT analysis, systematising information (expected outcomes from each intervention), designing an IL draft, and validating the IL with stakeholders.

- **Experts workshops**

The selection of indicators to evaluate the effectiveness of the interventions included in the IL was conducted through the implementation of workshops with other researchers, and a Delphi survey.

Results and practical innovations

The IL clarified **causal pathways between interventions, and impacts**, helping to identify more relevant indicators for the Irish context and current data gaps for the monitoring and evaluation of direct payment interventions. It highlighted where **existing metrics failed to capture outcomes** such as farm resilience, generational renewal, and agricultural diversification.

Funded by the European Union

Tools4CAP has received funding from the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement No. 101086311. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Practical innovations included the use of **participatory scoring** to prioritise indicators and the development of **an IL template that can be embedded into annual performance reviews**. This template has the potential to be used in other policy areas.

By visualising how payments translate into measurable outcomes, the IL strengthened both **internal management** and **external accountability mechanisms**.

Benefits, limitations, and risks

Benefits include enhanced clarity of intervention design, improved **alignment** between objectives and results, and better-informed **performance reporting**. The IL provides a **transparent reference framework** for policy communication and evaluation.

Although it was not possible to develop all the proposed indicators as sufficient data was lacking or **baseline values not in existence**, the case study identified data needs, guiding future collection requirements.

Nonetheless, the approach improved institutional knowledge and encouraged **regular reflection** on policy effectiveness.

A potential risk from the implementation of the tool is the **lack of active participation** of relevant stakeholders.

Recommendations

This tool has **potential to be implemented by other Member States** (MS). The IL process should be institutionalised by Ireland and other MS within its CAP monitoring structures, updating it annually as new data and evaluations emerge. Integration with the PMEF dashboard will allow for **continuous learning and adaptive management** in terms of policy effectiveness. From the perspective of the user, additional **resources for data management and staff training** are recommended to maintain consistency.

The approach should also be extended to **Pillar II interventions**, ensuring that rural development measures are evaluated with equal rigour.

Ultimately, embedding IL within the governance culture will strengthen the transparency and responsiveness of CAP policy in Ireland.

TOOLS4CAP

CASE STUDY FACTSHEET

Policy Monitoring Tool for Performance Assessment

ITALY

Description and relevance

The Italian case study tested the Policy Monitoring Tool (PMT) as an instrument for assessing CAP Strategic Plan performance through the integration of multiple datasets and stakeholder feedback. Conducted by ABACO, it aimed to evaluate the tool's ability to **support both ex-ante design and ex-post evaluation** by visualising results indicators in real time.

The initiative addressed **Italy's complex governance environment**, where numerous regional administrations implement CAP measures. The PMT was designed to **bridge information gaps**, provide a **harmonised analytical framework**, and **reduce administrative burdens** by automating data flows.

The case aligns with the New Delivery Model's emphasis on performance-based management, showcasing **how digital tools can translate diverse datasets** into actionable insights for policymakers.

ABOUT THE CASE STUDY

TOOL CATEGORY: **Data and monitoring systems**

GEOGRAPHIC LEVEL: **Regional**

TYPE: **Trial run + Showcase**

Christian Schingo, *CS coordinator*

Results and practical innovations

The Italian pilot successfully demonstrated PMT's functionality using synthetic datasets representing CAP interventions. **Interactive dashboards** visualised trends in **income support**, **competitiveness**, and **environmental performance**.

Stakeholder workshops validated the tool's usability and provided suggestions for enhancements, including integration of farm-level and regional data. The process highlighted the **importance of co-design between technical experts and policy users**.

Innovations included **simplified data pipelines**, improved **visual analytics**, and structured **feedback loops** between CREA and regional authorities. Despite being a prototype, PMT proved capable of harmonising monitoring across regions and **offering transparent**, **replicable methods for policy review**.

Funded by the European Union

Tools4CAP has received funding from the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement No. 101086311. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Tools applied

- **Policy Monitoring Tool** (PMT) prototype.
- Multi-source data integration and indicator **dashboards**.
- **Stakeholder workshops and feedback** questionnaires.
- Comparative **analysis** with CAP PMEF framework.

Benefits, limitations, and risks

Benefits include **improved accessibility** of performance data, **stronger coordination** among institutions, and **increased analytical capacity**. PMT supports faster reporting and **enhances accountability** by visualising results indicators dynamically.

However, challenges involved **inconsistent data availability**, **technical issues** during demonstrations, and **low response rates** to feedback surveys.

Risks relate to **fragmentation** between regional systems and limited institutional ownership if not embedded in official workflows. Yet, the case confirmed that **user-friendly digital tools can transform CAP governance** by enabling data-driven dialogue.

Recommendations

Italy should prioritise expanding PMT coverage to real administrative datasets, ensuring interoperability with regional systems. **Continued stakeholder training** and iterative improvements to usability will sustain engagement.

Embedding PMT into formal monitoring processes and connecting it with EU-level platforms would **enhance consistency**. Future development should include **real-time alerts for performance deviations** and **automated data validation**. By institutionalising PMT use, Italy can improve coordination across governance levels and strengthen the overall credibility of CAP performance assessment.

TOOLS4CAP

CASE STUDY FACTSHEET

Integrated Tools for Eco-Scheme Design and Monitoring

NETHERLANDS

Description and relevance

The Dutch case study integrated three complementary tools—the IOI matrix, the joint eco-scheme–farm income model, and the Policy Monitoring Tool—to **optimise the design, targeting, and monitoring of eco-schemes** under the CAP. The aim was to demonstrate how combining qualitative and quantitative approaches can **improve evidence-based policymaking**.

The case centred on the national eco-scheme using a **points-based system rewarding environmental performance**. It addressed the challenge of **balancing economic viability with ecological effectiveness**.

Wageningen University and Research led the study in collaboration with the Ministry of Agriculture, Fisheries, Food Security and Nature, testing how integrated tools could support decision-making throughout the policy cycle **—from prioritisation to evaluation**.

ABOUT THE CASE STUDY

TOOL CATEGORY: Integral tool (multi-tool approaches)

GEOGRAPHIC LEVEL: National

TYPE: Full adoption: Joint eco-scheme, Farmdyn, AGMEMOD farm income models**; Trial run:** IOI matrix**; Showcase:** Policy monitoring tool

WAGENINGEN UNIVERSITY & RESEARCH

Marcel van Asseldonk
Ana Gonzalez-Martinez
Roel Jongeneel
John Helming
Marit Tonkens

Results and practical innovations

The integrated framework allowed the Dutch team to assess eco-scheme design and performance comprehensively. The IOI matrix **clarified strategic priorities**, while the combined FARMDYN–AGMEMOD model linked farm-level and market-level impacts, **testing payment calibration and uptake scenarios**. The Policy Monitoring Tool explored **digital verification** using satellite and geotagged photos. Together, the tools produced a coherent picture of eco-scheme feasibility and trade-offs.

Innovations included **inter-model data exchange**, **participatory evaluation** with policymakers, and the **use of AMS data** for monitoring compliance. The integration showed how linking modelling, governance, and monitoring tools supports adaptive and results-oriented CAP implementation.

Funded by the European Union

Tools4CAP has received funding from the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement No. 101086311. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Tools applied

- **IOI Matrix** for needs assessment and instrument mapping.
- **FARMDYN and AGMEMOD** economic models.
- **Policy Monitoring Tool** (AMS/LPIS data integration).

Benefits, limitations, and risks

Benefits included **improved coordination** across analytical domains, **stronger evidence** for policy design, and **enhanced transparency**. The integrated approach **bridged economic modelling with practical monitoring systems**.

Limitations were **primarily institutional**—limited staff time, data heterogeneity, and political uncertainty. Risks involved potential **misinterpretation of model outputs** and **dependency on external expertise**.

Despite these challenges, the Netherlands demonstrated that combining tools can transform policymaking **from reactive to proactive**, enabling more strategic use of data and modelling.

Recommendations

To sustain progress, **integration should become standard practice in CAP planning**. Ministries should maintain a cross-disciplinary task force linking modellers, data managers, and policy officers. **Simplified communication materials** explaining complex results can improve stakeholder understanding. **Investments in interoperability** and **open data standards** are essential to maintain coherence between models and monitoring systems. Continued collaboration with EU partners will further refine this holistic, evidence-driven approach to CAP governance.

TOOLS4CAP

CASE STUDY FACTSHEET

Prioritisation Tool for Ranking CAP Needs

POLAND

Description and relevance

The Polish case study focused on testing a participatory Prioritisation Tool designed to enhance transparency and inclusiveness in **ranking the needs identified** for the next CAP Strategic Plan. The process aimed to **address long-standing challenges in policy design**, where limited time and uneven stakeholder participation often **constrained debate** and **reduced objectivity**.

The case built upon lessons from the 2023–2027 CSP cycle, where prioritisation was seen as **overly technocratic**. By introducing a **structured participatory approach**, the study sought to ensure balanced input from representatives of the **"quadruple helix"** – policymakers, practitioners, academia, and civil society.

The experiment was conducted by the European Rural Development Network (ERDN) team in collaboration with national stakeholders. It demonstrated how a **transparent voting mechanism** could **foster better understanding of trade-offs** among social, economic, and environmental priorities in agricultural and rural development policy.

ABOUT THE CASE STUDY

TOOL CATEGORY: **Qualitative and participatory methods**

GEOGRAPHIC LEVEL: **National**

TYPE: **Showcase**

Barbara Wieliczko, *CS coordinator*
Zbigniew Floriańczyk

Results and practical innovations

Workshops revealed that the **modified CV method** was simple enough for diverse participants while rigorous enough to provide meaningful results. Each participant distributed 100 points among needs, constrained by minimum and maximum thresholds to avoid dominance by specific interests. This ensured fairer representation of minority perspectives.

The process revealed **convergence around themes of sustainability, farm resilience, and knowledge transfer**, while exposing **tensions between production-oriented and environmental priorities**.

Innovations included **pre-testing** the method, **real-time feedback** on aggregated results, and **transparent documentation** of each stage. The exercise demonstrated how participatory prioritisation can both **democratise policymaking and increase stakeholder ownership** of CAP strategies.

Funded by the European Union

Tools4CAP has received funding from the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement No. 101086311. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Tools applied

- **Constrained Cumulative Voting** (CV) procedure.
- **Stakeholder workshops** with representatives of the quadruple helix.
- Facilitated discussion and **documentation protocol**.
- **Scoring tables** linking needs to CAP objectives.

Benefits, limitations, and risks

The principal benefit was the establishment of a **transparent and inclusive prioritisation process** that participants perceived as **legitimate and fair**. It improved dialogue among groups traditionally separated in policy debates and promoted **evidence-based reasoning**.

However, **time constraints** and **limited political influence** of the exercise restricted the scope of **stakeholder participation**.

Risks included potential **tokenism** if the process remains disconnected from actual decision-making. Moreover, **differences in technical literacy** among stakeholders occasionally created asymmetries in understanding the preferences of other stakeholders.

Nevertheless, the method successfully demonstrated that a structured voting mechanism can **enhance both objectivity and social legitimacy** of CAP planning.

Recommendations

Poland should institutionalise participatory prioritisation early in the CSP design process, ensuring real policy impact rather than symbolic consultation. **Regional replication** could test **scalability and reveal local variations** in preferences.

Continuous capacity building for facilitators and stakeholder representatives is crucial to maintain quality. In the long term, integrating the prioritisation tool with quantitative indicators from monitoring systems would **bridge stakeholder perceptions with measurable policy outcomes**, reinforcing the credibility and inclusiveness of CAP governance.

TOOLS4CAP

CASE STUDY FACTSHEET

Integrated Vision-Building Tool for Agricultural Policy 2040

SLOVENIA

Description and relevance

The Slovenian case study supported the Ministry of Agriculture, Forestry and Food (MAFF) in preparing a **participatory, evidence-based vision** for the future of food, agriculture and rural development by 2040. It aimed to build an **integrated process combining participatory foresight, scenario analysis,** and **indicator-based monitoring**.

The rationale was to provide a shared long-term policy framework, **aligning CAP interventions with broader sustainability and food system goals**.

The University of Ljubljana's Biotechnical Faculty steered the process in collaboration with MAFF, using expert group meetings and stakeholder consultations to define strategic objectives. The approach aimed for **inclusion, legitimacy, and data-driven dialogue** to ensure that the future vision reflected both scientific evidence and evolving societal demands, ultimately serving as a foundation for the next CAP Strategic Plan.

ABOUT THE CASE STUDY

TOOL CATEGORY: Integral tool (multi-tool approaches)

GEOGRAPHIC LEVEL: National

TYPE: Full adoption: Participatory Vision Building, Policy monitoring tool; **Ready-to-use:** Scenario building

UNIVERZA V LJUBLJANI
University of Ljubljana

Ilona Rac, *CS coordinator*
Emil Erjavec
Matic Soklič

Tools applied

- **Participatory Vision-Building**

A methodological approach designed to strengthen inclusiveness and transparency in policymaking by involving diverse stakeholders in a co-design process. In the Slovenian case, the method was applied in two steps: first, a draft vision was developed through expert consensus; second, broader deliberations with selected stakeholders refined and validated the vision. The agreed text is currently awaiting final validation at higher political levels (Government and Parliament).

- **Scenario Building for Impact Assessment**

The approach constructs alternative futures to explore different agricultural policy options. The method supports evidence-informed discussions of trade-offs and risks, improving strategic foresight and transparency in selecting policy directions.

Funded by the European Union

Tools4CAP has received funding from the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement No. 101086311. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

- **Indicator-based Policy Monitoring Tool**

The tool draws on indicator-based monitoring frameworks, in particular the CAP's Performance Monitoring and Evaluation Framework (PMEF) in combination with selected national statistics data. It provides a structured basis for both scenario analysis and vision building.

Results and practical innovations

The process produced a broad dataset of **around 120 sustainability indicators, trend analyses to 2040**, and **alternative policy scenarios** exploring economic, environmental, and social aspects of sustainability. Expert discussions and public consultations refined a draft vision centred on **agreed strategic objectives**.

Innovations included integrating multiple tools into a single policy process and using **a systems map** as a visual tool to capture and discuss interdependencies across the agri-food system. The approach entailed cross-sector collaboration and exposed **data gaps in a number of areas such as health, consumer behaviour, and food systems knowledge**.

The vision-building exercise became both a technical and institutional learning process, aiming to establish a foundation for long-term, evidence-based policymaking.

Benefits, limitations, and risks

Key benefits included enhanced **analytical capacity** within the MAFF, confrontation of different strands of highly specialised, siloed expert knowledge and stakeholder views, strengthened interdepartmental cooperation, and improved appreciation of policy trade-offs. The participatory element provided a chance to increase **legitimacy and ownership** among stakeholders, though the process is still ongoing.

Limitations emerged from **incomplete or missing datasets, delayed inclusion** of some stakeholders, occasional **institutional changes** and information asymmetries between decision-making levels. Furthermore, even though the process was supported by a relatively sound evidence base, **stakeholder discussions still often reverted to clichés and opinion-based declarations**, revealing a rather weak ability to relate data indicators to subjectively understood issues.

Risks also included potential **dilution of ambition** during political transitions and challenges in finalising the vision text. Finally, the political fate and subsequent legitimacy of the document are still uncertain.

Recommendations

Member States should consider **conducting regular fact-based and inclusive agenda-setting** processes with regular and formalised engagement **with academia and civil society** in relation to CAP Strategic Planning (or its equivalent in future programming periods) to promote a culture of co-creation, legitimacy, ownership and policy learning.

The weak ability to link data and complex substantive issues in the policy community revealed the need to **strengthen analytical capacity**. It is highly likely that the data and knowledge gaps revealed in the Slovenian vision-building process are also present in other countries.

The same can be stated about **limited indicator coverage for non-traditional issues** in agricultural policymaking. Food systems' aspects of social sustainability, consumer preferences and public health are all poorly covered in both national and EU databases, not only due to poor data coverage but because of **poor conceptualisation**, indicating a strong need to strengthen the indicator framework.