EVALUATION OF METHODS AND TOOLS

FOR THE DESIGN AND MONITORING OF THE CAP STRATEGIC PLANS

D 1.3

FEBRUARY 2024

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

EVALUATION OF METHODS AND TOOLS

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	Daniele Bertolozzi-Caredio (ECR), Francesco Iadecola (ECR), Bérénice Dupeux (ECR), Eleonora Santoni (ECR), Ralph Rustom (ECR), Laura Van den Bossche (ECR), Roel Jongeneel (WUR), Ana Gonzalez Martinez (WUR), Marcel van Asseldonk (WUR), Simone Sterly (IFLS), Carla Wember (IFLS), Nicolaos Kalatzis (NP), Vasilis Pyrgiotis (NP), Giorgos Charvalis (NP), Emil Erjavec (UL), Ilona Rac (UL), Blanca Casares (AEIDL), Miranda García Lera (AEIDL), Peter Witzke (EuroCARE), Guna Salputra (EuroCARE), Petr Havlik (IIASA), Christian Schingo (ABACO), Trevor Donnellan (TEAGASC), Emma Dillon (TEAGASC), Victor Kieffer (OB), Hannah Fried (OB), Zbigniew Floriańczyk (ERDN), Barbara Wieliczko (ERDN), David Nafría García (Itacyl), Alberto Gutiérrez García (Itacyl), Thomas Fellmann (JRC), Caetano Beber (JRC), Bozhidar Ivanov (IAE-BG), Zsolt Szabó (AKI), János Sávoly (AKI), Norbert Potori (AKI), Weronika Felcis (UoL), Živilė Gedminaitė-Raudonė (LSMC), Rita Lankauskiene (LSMC), Vitalija Simonaityte (LSMC), Monica Mihaela Tudor (IAE-RO), Camelia Gavrilescu (IAE-RO), Marieta Okenkova (SUAN), Zuzana Palkova (SUAN)
Work Package Leader	Ecorys Belgium
Project coordinator	Ecorys Belgium

Table of contents

List of figures

List of tables

Acronyms

AKIS	Agricultural Knowledge and Innovation System
AMS	Area Monitoring System
CAP	Common Agricultural Policy
CDE	Communication, Dissemination, Exploitation
CSP	CAP Strategic Plan
FADN	Farm Accountancy Data Network
EAFRD	European Agricultural Fund for Rural Development
EC	European Commission
EGD	European Green Deal
EU	European Union
GA	Grant Agreement
IACS	Integrated Administration and Control System
ICT	Information and Communications Technology
IT	Information Technology
JRC	Joint Research Centre
KPI	Key Performance Indicators
MCA	Multi-Criteria Analysis
MS	Member State
Ms	Milestone
NDM	New Delivery Model
NGO	Non-Governmental Organisation
OECD	Organisation for Economic Co-operation and Development
RDP	Rural Development Programme
SDG	Sustainable Development Goals
SEA	Strategic Environmental Assessment
SMART	Specific, Measurable, Achievable, Relevant, and Time-Bound
SO	Specific Objective
SWOT	Strengths Weaknesses Opportunity Threats
TEF	Tune-up Evaluation Framework
Tools4CAP	Innovative Toolbox empowering effective CAP governance towards EU ambitions
TWG	Thematic Working Group
UAA	Utilised Agricultural Area
VP	Voting and Prioritisation
WP	Work Package

1. Introduction

1.1. Why Tools4CAP

In the context of a changing Common Agricultural Policy (CAP), the New Delivery Model (NDM) established in the Regulation EU 2115/2021[1] introduces the CSPs and new monitoring, review and evaluation requirements. In 2019, the European Commission (EC) launched the European Green Deal (EGD) including the Biodiversity Strategy, the Farm to Fork strategy, the Soil and Forest Strategy and the Climate Adaptation plan and set up sustainability targets to achieve mostly by 2030, also through the CAP. The new CAP comes with a performance-oriented rather than compliance-based approach, more responsibilities and flexibility for MS for the design of the SPs and the monitoring framework, as well as a new policy cycle entailing exchange and coordination with the European Commission.

MS are called to acquire new capabilities and set up proper methodological tools to ensure i) alignment with the CAP, GD and SDGs objectives 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; and 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, therefore, enters into play in the context of a changing Common Agricultural Policy (CAP) and aims to provide CAP decision-makers with suitable tools for a more evidence-based policy making, ultimately improving capabilities to design next generation SPs, and to perform monitoring tasks. To realise its ambitions, the project pursues five specific objectives:

Objective 1. To provide a shared knowledge base and an evaluation of methods and tools used for the design and implementation of the SP.

Objective 2. To identify and adapt innovative methods and tools for the design and implementation of the SP, by taking stock of relevant and replicable solutions developed in recent and ongoing research projects and other EU initiatives.

Objective 3. To empower end users to adopt innovative solutions for the design and implementation of the SP, by providing them with methodological guidance on choosing the best solutions, their operationalisation, and associated good practices.

Objective 4. To establish a replication lab supporting the practical demonstration and uptake of innovative solutions for the design and implementation of the SP, by operationalising and testing methods and tools across case studies.

Objective 5. To set up a capacity building hub to mobilise knowledge and transfer operational capabilities to end users for the design and implementation of the SP, by enabling mutual learning, participation, and science-policy dialogue.

1.2. The objectives of the deliverable

The main goal of this deliverable is to evaluate methods and tools used by MS for the design and monitoring of their CSPs 2023-2027. The rationale behind the evaluation is to shed light on the strengths and weaknesses of the different tools employed, in order to highlight possible areas of improvement, identify in which tasks of the policy design and monitoring they can be most useful, and ultimately facilitate their replication for the next generation CSPs. The evaluation focuses on tools already employed by MSs for the CSPs 2023-2027; however, we also included policy analysis tools that are commonly employed by the European Commission for the evaluation of the CAP, as they could complement the activities carried out by MSs. It is also important to underline that the evaluated

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

tools are not the only options available to MSs for the next CSPs, but manifold tools are also available in science (or other policies, for example climate or environmental) that have not been implemented for the current CSPs, but may be of interest to MSs for the next generation CSPs. These tools, however, will be explored and assessed in future tasks of the Tools4CAP project.

Based on multiple sources of information, including interviews, online surveys and focus groups with national experts involved in the design and monitoring of the CSPs, but also literature review and the expert knowledge available within our consortium, the evaluation is performed against a set of five criteria and 12 sub-criteria. Evaluation questions are defined to address each sub-criterion. Lastly, benchmarking factsheets are produced to highlight strengths and weaknesses of different types of tools adopted in relation to different tasks of the policy design (e.g. needs assessment, intervention setting, stakeholder consultation, etc.) and monitoring (compliance checks and performance review).

The deliverable illustrates the approach and evaluation strategy, and includes a description of the data collection, as well as the evaluation criteria and evaluation questions adopted for the assessment (Chapter 2). Moreover, D1.3 presents the results of the evaluation for each category of tools against the identified criteria and sub-criteria, and summarises the main key findings for each category of tool (Chapter 3). Additionally, the deliverable presents benchmarking factsheets (in attachment) focused on specific tasks of the policy design and monitoring, which report on good practices that emerged during the evaluation.

The results of the evaluation and the identified good practices will support the subsequent project's tasks, in particular in relation to the identification of the most interesting tools to be adapted and tested in case studies to boost their replication and uptake across MSs. The results of the evaluation will be also instrumental for MSs' end users to gain a comprehensive understanding of the available options, and their pros and cons, for the next CSPs.

1.3. Key definitions

1.3.1. Design and monitoring of the CAP Strategic Plans

According to Regulation EU 2115/2021, MSs are required to design CSPs covering their national territory and ensuring consistency with regional settings, and to establish a monitoring, reporting and evaluation system. The governance of the SP, therefore, begins with their design and continues throughout their implementation. Design and implementation are inter-functional phases of the same process that complement one another.

We define '**design'** as the process involving all activities performed by MS, including the preparatory work, aiming at drafting the CAP SP, hence at creating a "sound intervention logic" of the CSP and the intervention strategy referred to in Article 107(1), point (b) for each specific objective. The CSPs' design involves specific tasks, including socio-economic and context analysis, SWOT analysis, needs assessment, interventions setting, targets setting, financial allocation, ex-ante analysis and Strategic Environmental Assessment (SEA), and stakeholder consultations.

We define '**implementation'** as the process involving all activities performed by the EC (and other EU agencies) and MSs, taking place after the formal approval of the CSP (1st January 2023), aimed at monitoring and evaluating the adoption of planned interventions and budget, and the achievements of the targets and objectives. We distinguish monitoring activities from evaluation activities. **Monitoring** includes all activities performed by the EC and MSs to collect and aggregate data to measure CSPs' performance and check beneficiaries' compliance. **Evaluation** includes all activities performed by the EC and MSs to evaluate progress ex-post in CSP implementation and in achieving the CSP set out objectives and the environmental and climate-related commitments of the EU.

The **project fully covers the design phase of the plans and the monitoring process**. The ex-post evaluation of the CSPs is out of scope. Further definitions and an understanding of the key phases of the New Delivery Model are provided in Deliverable 1.2 – Tools4CAP's conceptual framework[2].

[2] https://www.tools4cap.eu/publications/

1.3.2. Methods and tools for CAP design and monitoring

The project focuses on the tools used for the design and monitoring of the CSPs. The term 'tools' encompass all methodologies, methods and (technical/technological) tools that can be used through the design and monitoring processes. Within an overarching methodological framework, methods are about instructions, procedures and processes for attaining a certain objective, alongside which different qualitative and quantitative tools are used. A tool is a device/product that help in accomplishing a task. Good practices will focus on the right method to use the different tools (e.g. method for stakeholder engagement or intervention logic). Focusing on the end product, the use of the tool can be appropriate or not (e.g. models, stakeholder platforms, etc.). For simplification, hereinafter, we refer to methods and tools exclusively with the term "tools".

1.3.3. End users

In order to foster the adoption of innovative tools, Tools4CAP aims to engage tools' end users to provide them with methodological guidance and capacity building opportunities. All actors playing a role in the CAP design and monitoring processes at either the regional, national or EU level are considered the end users of the Tools4CAP coordination and support action. For both the design and monitoring phases, end users operate (or play a role) at different levels of the governance. The CSPs, for instance, identify a mix of actors including control bodies, managing authorities, paying agencies, competent authorities, certification bodies, monitoring committees, coordination and communication bodies, including AKIS. End users can also be research organisations, such as universities of research centres, and private organisations, such as consulting companies or IT developers.

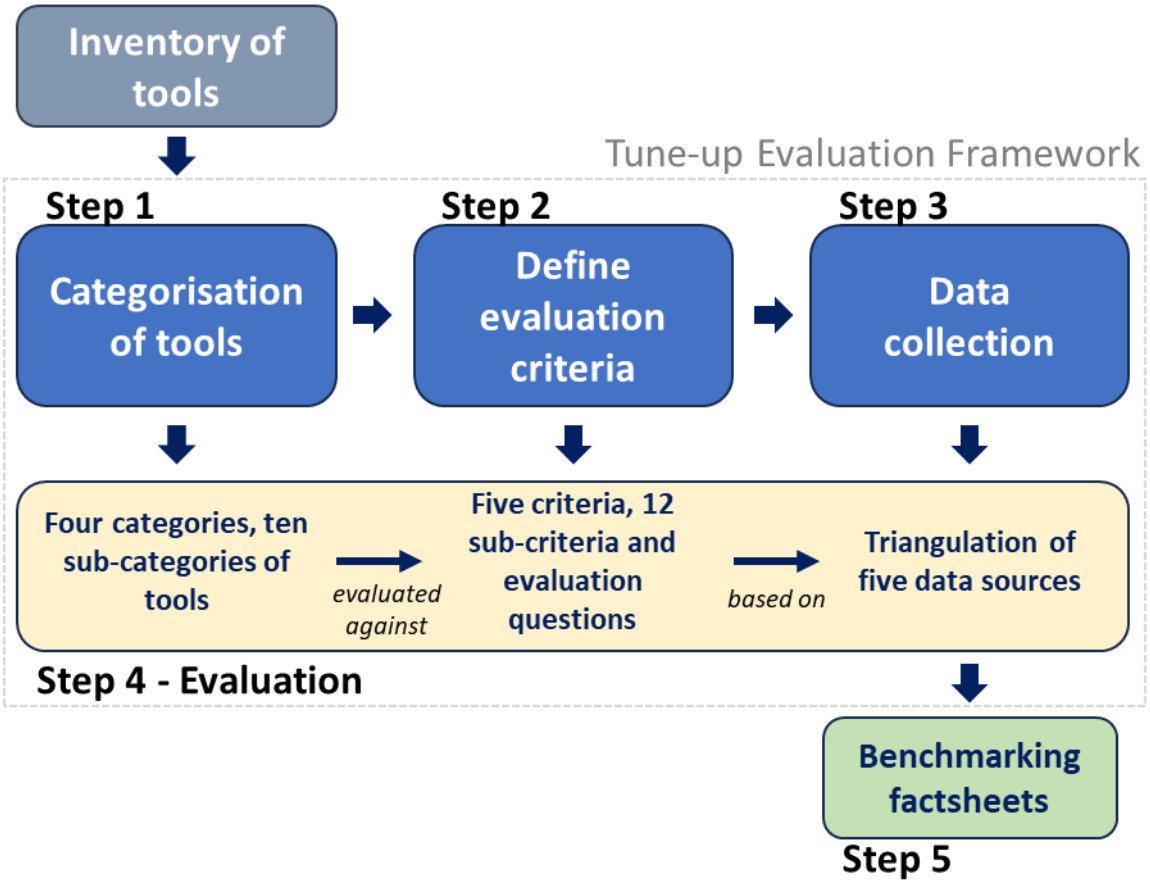
2. Methodology

2.1. General approach

The scope of the evaluation is limited to the tools identified in the inventory of tools built under task 1.1 of the project[3] and tools commonly used by the European Commission. These include for example policy analysis tools like CAPRI or AGMEMOD, or policy choices supporting tools outlined in the EU Better Regulation Toolbox, such as Multicriteria and Cosy-Benefit Analyses. Different types of tools are evaluated against a set of criteria and sub-criteria by triangulating information gained through different sources. The results of the evaluation shed light on the potential and limitations of adopting different types of tools, as well as practical aspects to account for when implementing such tools. Based on these results, benchmarking factsheets were produced, each one focusing on a different task of the CSP design or monitoring, outlining key strengths and weaknesses of the main tools adopted to perform the corresponding task. The evaluation results and the benchmarking factsheets, therefore, can help end users in the MSs account for the different pros and cons, as well as good practices, when adopting different tools. As such, they can also help guide the choice of the tools. As shown in Figure 1, our approach followed five steps. The steps from 1 to 4 constitutes the Tune-up Evaluation Framework (TEF), which is intended to provide a conceptual and methodological instrument for the analysis of tools specifically employed for the design and monitoring of the CSPs (the TEF is described in detail under section 2.2).

[3] https://www.tools4cap.eu/tools/

Figure 1. General approach of the evaluation.



In the first step, based on the inventory of tools, a categorisation of tools is produced. Four categories of tools were defined, further divided in 10 sub-categories of tools. The categorisation was carried out based on the methodological and technical characteristics of the tools, regardless of their application purpose (i.e. the different design or monitoring tasks for which they can be employed). The evaluation was performed for each sub-category of tool separately. Categorisation facilitates fair comparisons among tools/methodologies with similar characteristics. It does not prioritise one over another but aids in an informed choice of tools.

In the second step, the evaluation framework was defined, consisting of five criteria further divided in 12 sub-criteria. To facilitate the evaluation of the sub-criteria, one evaluation question has been formulated for each sub-criterion. The evaluation, by means of a triangulation of different information sources, seeks to answer the evaluation questions. The answers to the evaluation questions under each criterion provide an overall judgment of the sub-category of tools against that specific criterion.

In the third step, data collection activities were conducted. Five different information sources were included: interviews, an online survey, and national focus groups with end users in the MSs, extensive literature review, and the expert knowledge available within our consortium.

In the fourth step, we performed the evaluation at the level of each sub-category of tools (as defined in step 1) by answering the evaluation questions (as established in step 2), through the triangulation of the different information sources (as identified in step 3). Based on the answers to the evaluation questions, for each sub-category of tools, an overall qualitative judgment is formulated according to a quality scale.

In the fifth and last step, based on the evaluation results, benchmarking factsheets were produced to showcase good practices emerged during the evaluation. Each benchmarking factsheet focuses on a key policy task of the CSP design or monitoring, and identifies two or three specific tools used by MSs to conduct the corresponding task. The factsheet outlines key strengths and weaknesses of such tools, and shows the judgment of the five criteria for

each tool. The factsheets, therefore, showcase examples of tools used by other MSs (or the EC) to conduct specific policy tasks, and highlight the pros and cons to be accounted for. They also help narrow the lens down to specific tools.

2.2. Tune-up Evaluation framework

As mentioned above, our evaluation is operationalised through the TEF, which includes steps from 1 to 4 of the general approach. The TEF is intended to provide a conceptual and methodological instrument for the analysis of tools 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 and sub-categories of tools evaluated, as well as the criteria, sub-criteria and related evaluation questions used to guide the evaluation. Thus, for each sub-category of tools, 12 pre-defined evaluation questions were answered, whereby 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 specify that, while the definition of the criteria and sub-criteria was the same for all sub-categories of tools evaluated, the evaluation questions slightly differ among sub-categories of tools. 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 the sub-category of tools. For example, granularity and accuracy are characterised differently among quantitative modelling tools and participatory tools. However, the evaluation questions are always meant to assess the criteria, which overall concept remains the same for all tools. Below, we describe the categorisation of tools, and the definition of the evaluation matrix, whereas the data collection activities are described in the next section.

Table 1. Tune-up Evaluation Framework.

				Categories of tools									
			Category ->	Stakeholder needs assessment tools			Policy choices supporting tools		Policy analysis tools		Monitoring and data stocktaking tools		
			Sub-category ->	Online consultation and surveys	Workshops and conferences	Focus groups and meetings	Voting and prioritisation tools	Multicriteria analyses	Statistical analyses	Mathematical Programming Tools	Compliance monitoring tools	Performance monitoring tools	Data stocktaking tools
	Criteria	Sub-criteria	Evaluation Question (EQ)										
Criteria	Accuracy	Data accuracy and validity	EQ1										
		Tool representativeness	EQ2										
	Reliabiliy	Replicability	EQ3										
		Transparency	EQ4										
	Applicability	Multi-functionality	EQ5										
		Experience	EQ6										
		Granularity	EQ7										
	Accessibility	Ease of use	EQ8										
		Tool availability	EQ9										
	Efficiency	Time efficiency	EQ10										
		Cost-effectiveness	EQ11										
		Human resources efficiency	EQ12										

2.2.1. Categories and sub-categories of tools

Tools were grouped in categories and sub-categories. The first purpose of categorisation is to organise similar tools into groups based on shared characteristics and methodological purpose. This enables fair and systematic comparisons by providing a structured framework for identifying strengths and weaknesses. Moreover, while in principle single specific tools could be evaluated, many of them share common or very similar methodological and technical characteristics, whereby an analysis of specific tools would produce redundant and overlapping results. The categorisation was conducted on two levels. First, tools were grouped into four main categories according to their methodological purpose, i.e., whether they are used for assessing stakeholder needs, to support complex policy choices, to perform ex-post or ex-ante analyses, or to collect data. Secondly, within each category of tools, tools were grouped in sub-categories according to their methodological and technical characteristics. The categorisation has been a work-in-progress procedure, meaning that it has been refined multiple times while deepening into the content and analysis of the tools. The definition of the categories and sub-categories of tools is provided below.

Stakeholder needs assessment tools: this category of tools is based on qualitative methodologies, including but not limited to participatory approaches, that enable the identification and assessment of stakeholders' perspectives and needs, which can be used to inform policy analyses and policy choices. Qualitative tools add value by offering rich information on complex problems, capturing singular or unforeseen possible impacts, shedding light on how individuals with diverse interests and roles perceive and make sense of policies, amplifying the voices of often marginalised perspectives, initiating preliminary investigations to formulate policies, evaluate possible scenarios, and progressing towards comprehensive explanations.

These qualitative techniques can be employed alongside more quantitative methods, either sequentially or concurrently. They are pivotal in ensuring the transparency, inclusivity and legitimacy of CSPs and in aligning the CAP with the needs and aspirations of those directly impacted by agricultural policies. These tools are also important in bringing groups that may have more limited professional analytical capacity into the decision-making process (e.g. small-scale farmer unions). This category is further divided into three sub-categories of tools, as follows:

- **Online consultation and surveys**: Online consultations and surveys are tools that leverage digital platforms to gather feedback and insights from a broader audience and raise specific questions, including (a) consultation, which involves actively seeking the opinions of interested and affected groups and (b) participation, which is the active involvement of interest groups in the formulation of policies and approaches, or in the drafting of regulatory texts (OECD 2001). Online surveys specifically are tools used to collect qualitative and quantitative data from different groups.

- **Workshops and conferences**: these are in-person or virtual events that are designed for collaborative discussions, brainstorming, or decision-making and offer the possibility of including perspectives, opinions of a large group of people. Compared to online consultations and surveys, face-to-face interactions open up potential for active deliberation and co-design of policy processes. Including participation of stakeholders, experts and policymakers in events around discussion, brainstorming and decision-making increases collaboration and helps to better understand the problem, issues and risks, and to craft solutions that are more likely to meet users' needs and achieve other policy objectives. Face-to-face participation in policy-making can improve legitimacy and impact.

- **Focus groups and meetings**: these are tools designed for gathering insights from specific groups of participants in a controlled environment. This can include stakeholders, policymakers and experts. A focus group is a facilitated group discussion focused on a specific content, in which the discussion is moderated, oftentimes on the basis of a discussion guide after an introductory stimulus.

Policy choices supporting tools: these tools rely on logic-based methodologies to facilitate decision-making. The tools are particularly useful when dealing with complex systems (i.e. characterised by interdependencies, trade-offs, relationships, and interactions between their parts), and when multiple policy options are available, information comes from multiple sources and several actors are involved in decision-making. These tools help converge to a shared, coherent, consistent and logical policy choice. This category is further divided into two sub-categories of tools, as follows:

- **Voting and prioritisation tools**: refer to a spectrum of instruments and methodologies, such as number voting, cumulative voting, or hierarchical prioritisation, that employ voting or analysis mechanisms 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. They rely on a structured approach where individuals or groups assign scores or numbers to different choices based on their perceived importance, urgency, relevance, or other criteria. Voting and prioritisation tools are often complemented with participatory activities, and can help foster a collaborative environment in which every stakeholder has an opportunity to voice their opinions and preferences, thereby ensuring that the collective perspective is taken into account in the decision-making process.

- **Multicriteria analyses**: Multicriteria analysis (MCA) tools, also known as multicriteria decision analysis tools (MCDA), are systematic approaches and methodologies used to evaluate and compare alternatives based on multiple and often conflicting criteria or objectives. De Luca et al. (2017) describe them as an umbrella term for methods that enable the consideration of multiple conflicting criteria to reach rational, justifiable and explainable decisions. They can also provide an objective means to include participative methods in decision processes. These tools are designed to assist decision-makers in situations where complex and diverse factors need to be considered simultaneously.

Policy analysis tools for evidence-based decisions: These tools serve for generating (scientific or empirical) evidence through the analysis of policies, either ex-ante or ex-post, to inform decision-making, hence underpinning evidence-based policymaking. Tools for both ex-ante analysis (employed before policy implementation to evaluate the likely outcomes, benefits and risks associated with different policy options), and ex-post analysis (evaluation and comparison of past experiences, or similar experience elsewhere, as well as outcomes under different policies and policy mix analysis), can be employed along the design and monitoring process. This category is further divided into three sub-categories of tools, as follows:

- **Statistical analyses**: statistical methods can be considered as tools that utilise statistical techniques to derive causal relationships between policy and potential impacts. These include econometrics, which uses economic theory, mathematics, and statistical inference to quantify economic phenomena. In other words, it turns theoretical economic models into useful tools for economic policymaking.

- **Simulation models tools**: these are quantitative modelling tools that primarily focus on simulating scenarios, interventions, or impacts. They can be used to predict the outcomes of various policy decisions before they are implemented. As defined by the EU Better Regulation Toolbox (2023), models are stylised representations of the real world that are used to make projections or to assess the behaviour of a system under specific (policy) assumptions.

Monitoring and data collection tools: These tools serve to collect and make available (but not to interpret) the necessary information and data for the performance review of the CSPs, beneficiaries' compliance checks, and to inform policy analyses and policy choices. These tools allow for collecting different types of qualitative and quantitative data, information and knowledge, with different levels of accuracy and based on different sources. They can rely on different methodological, technical or technological tools. This category is further divided into three sub-categories of tools, as follows:

- **Compliance monitoring tools**: CAP compliance monitoring tools encompass a range of quantitative and qualitative instruments. These tools include specialised systems that rely on various technologies for collecting, storing and sharing data (e.g. satellite imagery, geographic information systems (GIS), and mobile applications). These systems could also offer processing capabilities providing the adequate information to ensure that beneficiaries comply with the conditionality requirements and eligibility criteria set for specific CAP interventions.

- **Performance monitoring tools**: Performance monitoring tools are integral for realising result indicators by providing an indication as to whether MSs are progressing towards or on track to reach the targets

established in their CSP. These tools are used in line with the Commission's prescribed methodologies around the use of specific data to calculate the result indicators[4]. These tools are developed for the regular gathering, analysis, and reporting of this data, and can involve both gathering and synthesising multiple datasets to allow the calculation of result indicators across different spatial and temporal scales, ensuring alignment with CAP objectives.

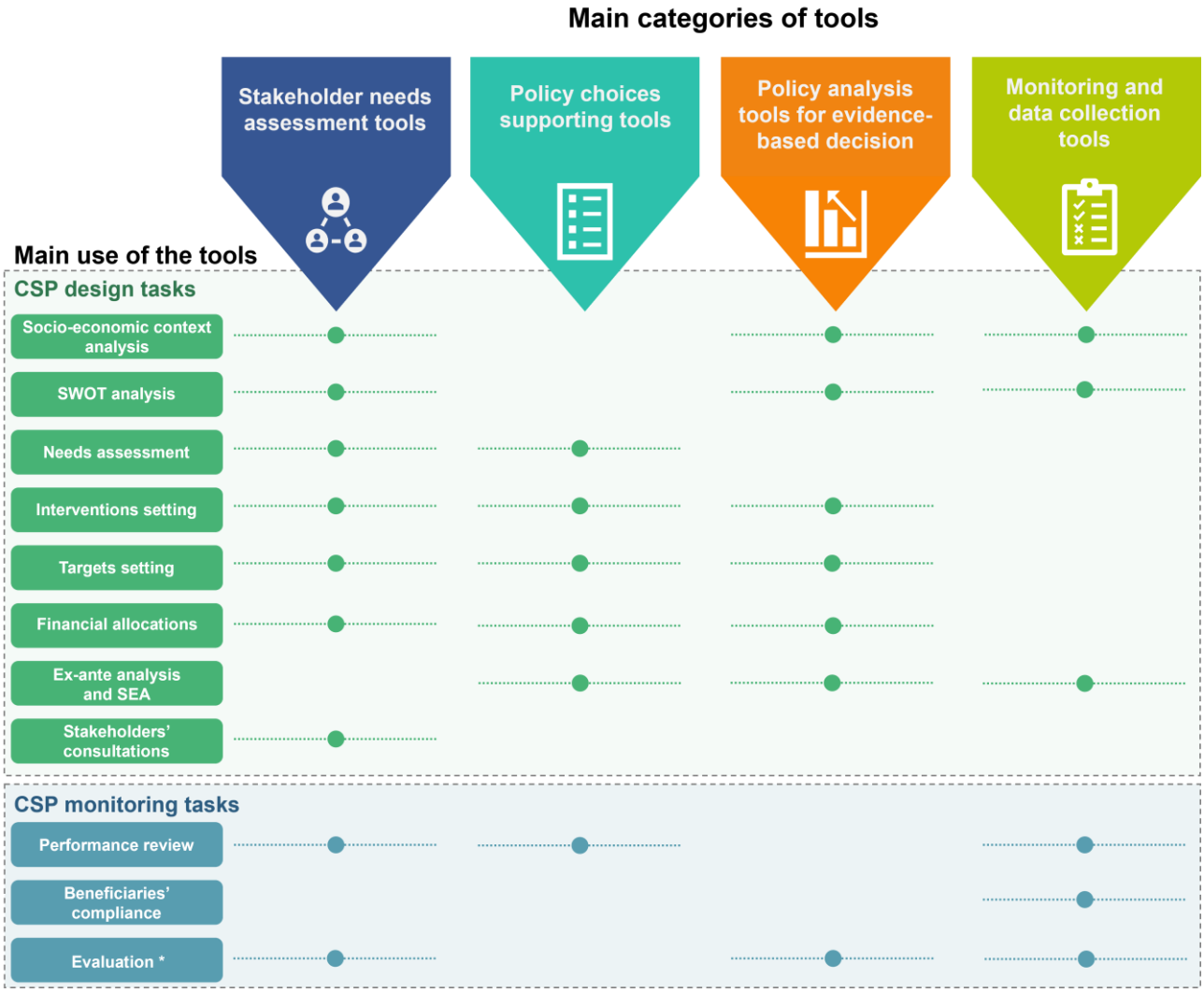
- **Data and knowledge stocktaking tools**: Monitoring and data collection tools for data and knowledge stocktaking facilitate the collection, storage and management of data and information necessary for policy monitoring and evaluation, informing policy choices. Some of these tools primarily serve other purposes, for example many are compliance or performance monitoring tools (e.g. related to FADN or IACS), but they also act as a data pool. These tools can generate data to construct various indicators, such as impact, context, and result indicators, which provide a more nuanced and comprehensive perspective on policy effects, playing a pivotal role in assessing and further informing policy strategies. They also play a key role in providing both the qualitative and quantitative data needed to compile annual performance reports and to ensure farmers' compliance checks. These monitoring and data collection tools also complement other policy analysis and policy choice supporting tools, providing essential data and insights.

In practice, diverse tools can be used for the same purpose or in combinations. For instances, tools can have multiple uses along the design and monitoring processes of the CSPs, whereby each design or monitoring task can be carried out through different tools. Secondly, tools can be complementary to one another, whereby each design or monitoring task can be carried out through a combination of tools, and the outputs of different tools correspond to the final choices making up a CSP fit for purpose. Figure 2 shows how the different categories of tools have been used or, in the case of monitoring tasks, are going to be used by MSs, based on the results of the Tools4CAP online inventory of tools[5].

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

[5] https://www.tools4cap.eu/tools/

Figure 2. Main categories and use of different categories of tools across design and monitoring tasks.

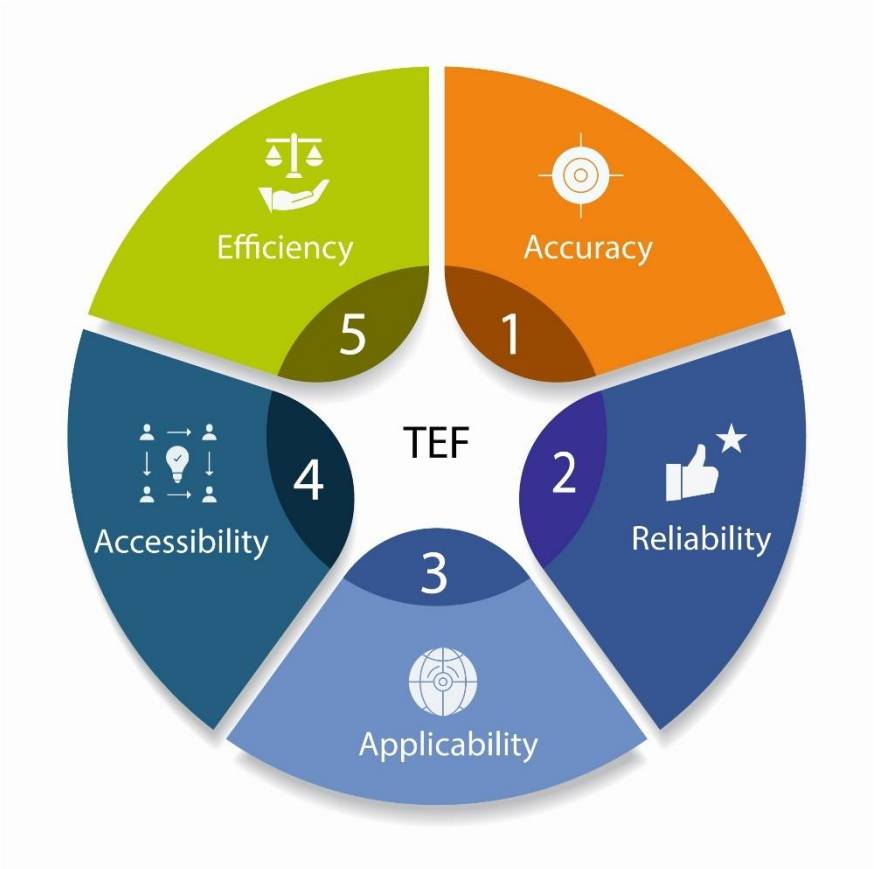


Note that the links between categories of tools and design and monitoring tasks were made based on the results of the Tools4CAP online inventory of tools (https://www.tools4cap.eu/tools/).
*Evaluation tasks are out of the scope of Tools4CAP. However, it is important to consider that different tools can also be used for evaluation.

2.2.2. Criteria, sub-criteria and evaluation questions

The TEF is based on the use of criteria and sub-criteria. This section describes the proposed criteria, sub-criteria, and the related evaluation questions. As shown in Figure 3, a total of five criteria have been defined to structure the analysis and ensure its comprehensiveness. The selection was made to draw operational conclusions on the technical and methodological aspects that can enable or hinder the implementation of the tool in practice. The set of criteria provides a strong basis for identifying proper evaluation questions aimed at investigating tools' characteristics or factors that contribute to make their application successful. 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 the tools. The five criteria are defined below.

Figure 3 – Tune-up Evaluation Framework criteria.



Accuracy. It reflects how successful the tool is in offering results that reflect reality without any distortion or omission. The evaluation of this criterion is important to make sure that end-users adopt tools that are capable of providing accurate results. Results that are distorted or omit critical information can lead to misinformed decisions based on flawed assumptions with serious consequences (e.g. waste of time and resources, missed opportunities, increased exposure to risks, stakeholders' discontent, damages to reputation, failure of public policies).

Reliability. It measures the degree of consistency and stability of the results obtained in the implementation process of the tool. This criterion measures the extent to which the process of implementation of the tool produces consistent findings when repeated. A tool that generates inconsistent results can be affected by non-systematic bias that hinders the validity and generalisation of the results.

Applicability. It 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. It 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 understand 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.

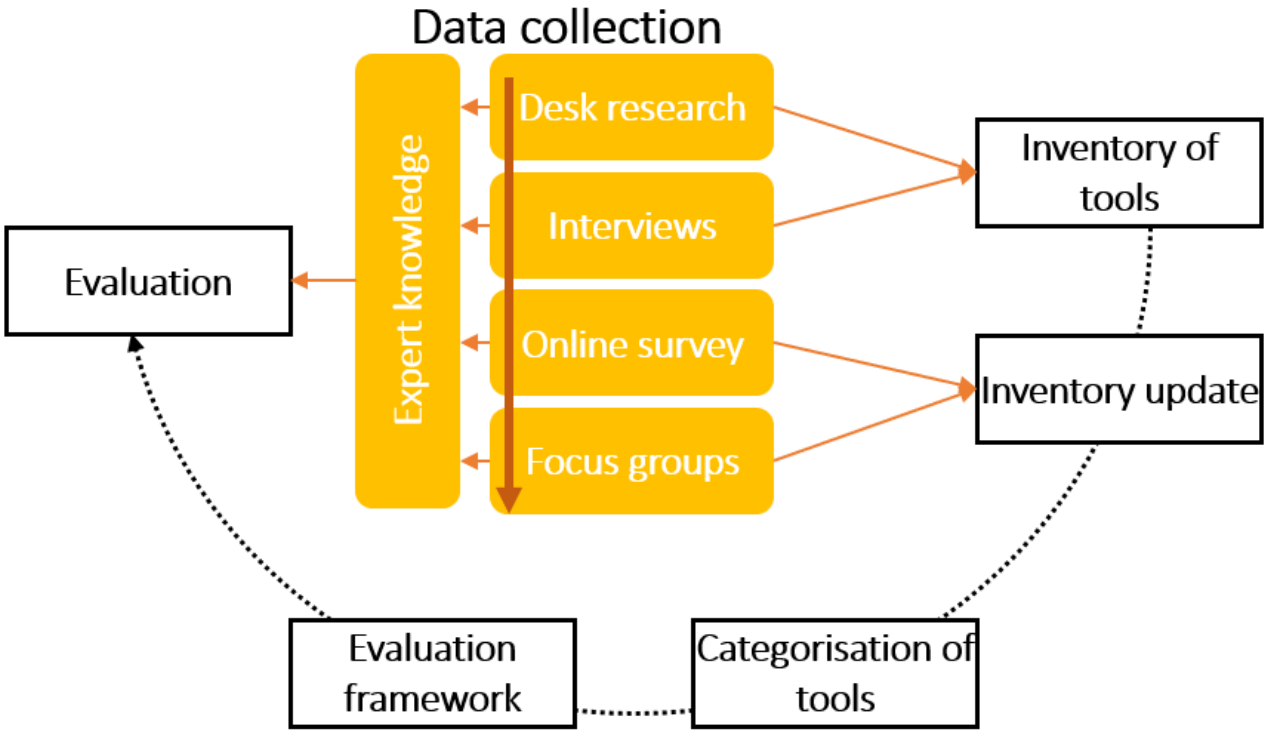
Efficiency. It measures the extent to which the tool is able to achieve desired outcomes given the resources available, including time, human resources, and costs. Moreover, this criterion is used to understand whether some tools achieved similar benefits (or greater benefits) at the same/lower cost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

As previously shown in Table 1, for each criterion, a set of sub-criteria has been developed. The sub-criteria act as building blocks that contribute to a nuanced and granular understanding of the subject matter. Each sub-criterion reflects a factor contributing to making up the criterion. Moreover, they serve as the basis for formulating evaluation questions guiding the analysis. One evaluation question was formulated for each sub-criterion. These questions provide a structured approach to guide the evaluation, ensuring that no critical aspect is left unexplored. As mentioned above, while criteria and sub-criteria are the same across sub-categories of tools, the evaluation questions differ slightly among sub-categories of tools. This allows to capture the different methodological and technical aspects making different types of tools satisfy a given criterion. For example, the methodological aspects making a participatory tool accurate are different from those that makes a quantitative model accurate. The definitions of the sub-criteria and the list of evaluation questions, are reported in Appendix B.

2.3. Data collection

The data collection strategy relied on a diverse array of sources, implemented at different stages of the project implementation to develop a comprehensive understanding of the tools under investigation. In particular, as shown in Figure 4, the evaluation relied on five data sources: semi-structured interviews, an online survey, 14 national focus groups, desk research, and consortium expert knowledge. The different data collection activities were conducted in sequential steps to allow one data source to complement the prior, and inform the next. Desk research and interviews were used to build the first draft inventory of tools, which, however, was affected by gaps. Online surveys and focus groups were used to fill the gaps and update the inventory. The consortium expert knowledge allowed to assess the collected information (consortium researchers were responsible for the analyses in their area of expertise), and provide further information to fill the gaps. All data sources were used in the evaluation. Each data collection activity is described below.

Figure 4. Diagram of data collection sources and their use.



2.3.1. Desk research

The desk research undertaken involved an extensive review and analysis of existing information from various public sources. This not only included scientific and grey literature, but also data from institutional websites, digital resources, and outputs from prior projects related to the CAP or EU agriculture. Importantly, the literature review also covered sources in national language across the EU-27, because important technical and dissemination information is often available only at national level. A significant part of this process involved examining academic

journals and government publications in local languages, providing literature, reports, and articles illuminating the tools used by EU MSs in the design and monitoring of their CSPs. Beyond traditional sources, the research also leveraged the vast resources of online databases and websites, which offer extensive collections of digital content, such as detailed reports summarising the various phases of the CSP process, as well as associated presentations, slides and press releases. The literature review was conducted in two steps. A first round of literature review was conducted at the beginning of the project to gather a basic understanding and preliminary mapping of the tools adopted across MSs. This step also supported the design of interviews and surveys. A second round of review was conducted after the interviews and surveys were collected, in order to enrich the collected information and fill the gaps.

This comprehensive and systematic approach to desk research allowed for an in-depth review of existing information, allowing the project team to trace out the activities undertaken at the various stages of the design and monitoring process. Moreover, the desk research allowed to identify additional methods and tools adopted for the CSP design and monitoring, as well as gathering information on their characteristics and implementation during different stages of the policy cycle. In particular, the review of scientific literature provides an in-depth understanding of the methodological aspects of the identified methods and tools, and allows to explore their use also in other fields (e.g. academic research) and contexts (e.g. adoption in other countries and/or for the design and monitoring of policies for other sectors).

2.3.2. Interviews

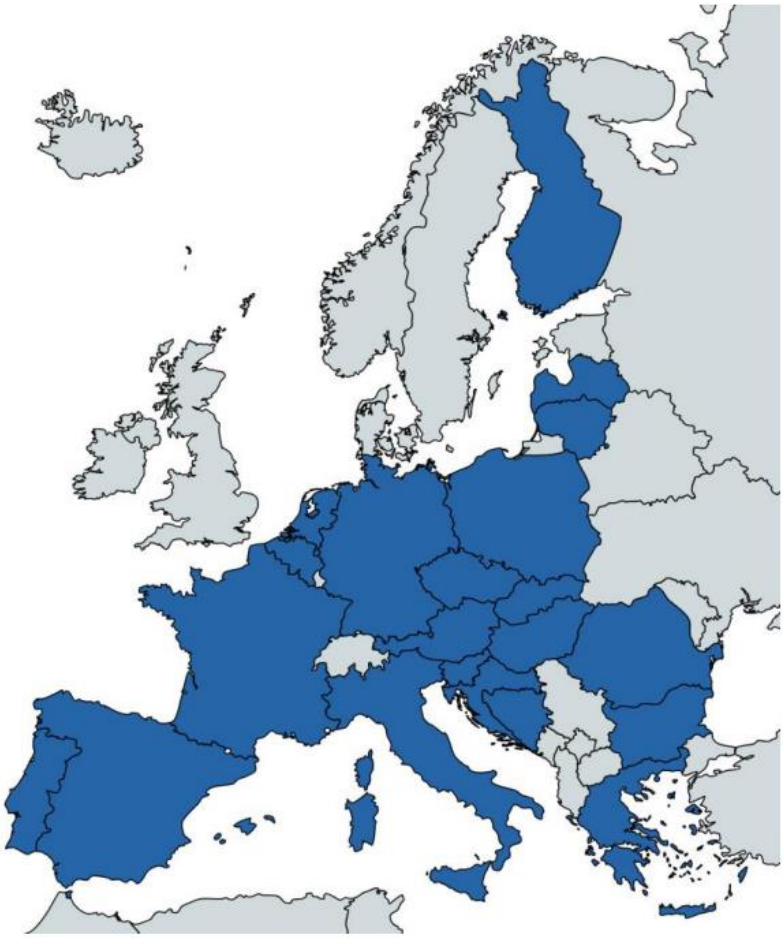
A total of 121 semi-structured interviews were conducted with end users across the 25 EU MSs. The remaining two MSs (Denmark and Estonia) were unable to be covered. In this case, extensive desk research was conducted to collect the necessary information. The interviews served for gaining an understanding of how each MS had undertaken the design and monitoring of the CSPs, as well as to identify and describe the main tools used by MSs for the design and monitoring. The interviews were conducted with stakeholders from various sectors, including ministries, payment agencies, other government entities, regional and local authorities, scientific and research institutions, consulting firms, farmer and agricultural organisations, as well as environmental and consumer groups across the EU. These interviews were carried out with end users operating at the local, regional, or national levels. These end users, whether engaged in past or ongoing decision-making processes related to CSP development or possessing substantial knowledge and experience in these domains, offered both detailed and comprehensive insights, including information not publicly accessible.

The interviews were conducted through a variety of means, including face-to-face meetings, video calls, and telephone conversations, allowing for flexibility based on the preferences and convenience of the participants. This approach ensured a broad spectrum of perspectives while preserving the personal, one-on-one nature of the interviews. We conducted the interviews in the national language of each MS. Figure 5 shows the number of interviews collected for each country. A detailed description of the collected interviews and the followed approach is provided in Deliverable 1.1[6].

[6] https://www.tools4cap.eu/

Figure 5. Collected interviews by country.

Country	Interviews
Austria	3
Belgium	6
Bulgaria	2
Croatia	4
Cyprus	4
Czechia	7
Denmark	0
Estonia	0
Finland	5
France	6
Germany	5
Greece	4
Hungary	5
Ireland	1
Italy	13
Latvia	5
Lithuania	8
Luxembourg	3
Malta	3
Netherlands	1
Poland	5
Portugal	8
Romania	6
Slovakia	4
Slovenia	6
Spain	2
Sweden	5



2.3.3. Survey

The interviews allowed for mapping the identified tools and for gathering a preliminary understanding of the factors at play in the choice of the tools and the tools' characteristics. Following the interviews, an online survey was launched with the aim to collect additional methodological and technical information on the identified tools The consultation was launched on EUSurvey and lasted between August and early September 2023. Like the interviews, the online survey was addressed to end users across the 27 MSs. The survey was translated into all EU languages to facilitate participants' responses.

The survey was tailored for each category of tools, ensuring relevance and efficiency in gathering specific insights. This led to the design of four different questionnaires, each focusing on one category of tools. The questions differed across the four questionnaires, because the four categories of tools are characterised by different technical and methodological aspects. Respondents were allowed to fill in one or more of these questionnaires of their choice, depending on their knowledge and expertise. The questionnaire contained both open-ended and close-ended questions, and collected both qualitative and quantitative information. The questionnaires are reported in Appendix A.

The consultation resulted in 77 valid contributions[7] from the stakeholders, distributed as presented in the table below (Table x). The survey gathered insights from relevant stakeholders from 16 MSs. The most represented stakeholder group in terms of respondents to the survey is "Scientific and research institutes", representing 43%, followed by

[7] A total of 76 responses gathered: 1 contribution considered not valid because the respondent did not include the name of the method/tool.

"Ministries" with 21% of responses, and "Consulting firms" (12%). Table 2 shows the number of contributions collected for each category of tools.

Table 2. Online surveys collected by category of tools

Table 1	Number of contributions
Stakeholder needs assessment tools	39
Policy choices supporting tools	16
Policy analysis tools for evidence-based decisions	8
Monitoring and data collection tools	14

2.3.4. Focus groups

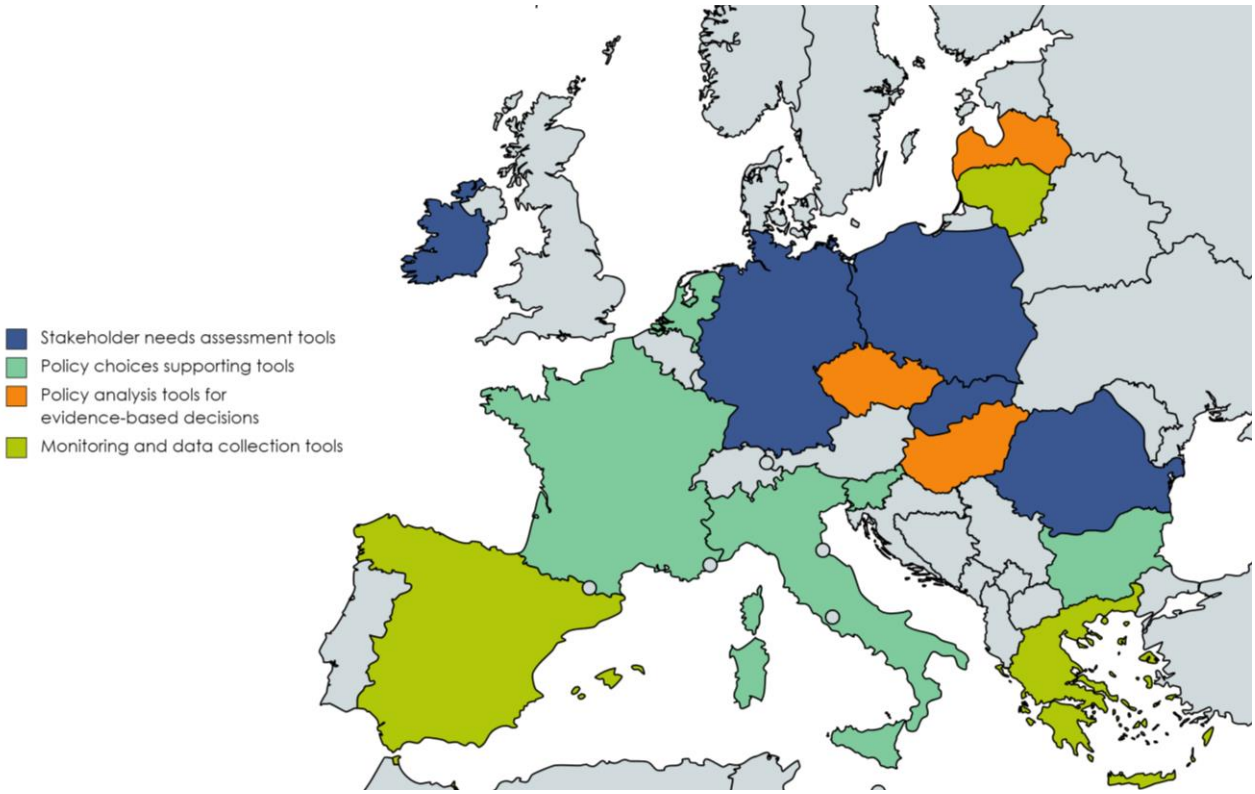
A total of 14 national focus groups were conducted in as many MSs. The aims of the focus groups were to (i) complement the inventory of tools (i.e. identify further tools that were not capture through the interviews), (ii) describe in detail the functioning of the tools in relation to the policy tasks for which they were employed, and (iii) identify key challenges and limitations encountered. The focus groups involved a diverse range of participants, including end users, policymakers, NGOs, agricultural representatives, environmental advocates, agricultural experts, and rural stakeholders. Each focus groups focused on a single category of tools, in order to allow for gathering in-depth information on specific tools.

The focus groups were conducted by national researchers within the consortium. The researchers were instructed on the objective and approach through multiple rounds of training. A standardised protocol was followed by all focus groups' researchers to report the collected information. However, room was left to national researchers to tailor the participants list, organisation and discussion within the focus groups according to their needs. As a standard rule, all focus groups lasted a minimum of three hours, and involved at least end users. However, some focus groups involved other types of stakeholders, whose perspective was considered useful for the scope of the focus groups.

Figure 6 shows the 16 MSs where the focus groups were located and identifies on which category of tools the focus groups were focused. Five MSs focused on stakeholder needs assessment tools, five MSs covered policy choices supporting tools, three MSs policy analysis tools, and three MSs monitoring tools. However, focus groups in Greece and Czechia have not been conducted yet, thus their output is not included in this evaluation. A total of 116 participants were involved in the 14 performed focus groups, of which 68 male and 48 female. The methodology and results of the national focus groups are reported in detail in Deliverable 6.5[8].

[8] https://www.tools4cap.eu/

Figure 6. Location and category of tools discussed in the 16 national focus groups.



Note: focus groups in Greece and Czechia have not been conducted at the time this report was conducted.

2.3.5. Expert knowledge

The Tools4CAP consortium gathers experts on the different categories of tools. As a final step, and based on the information collected through the desk research, interviews, online survey and focus groups, the consortium experts provided their input to enrich, validate and consolidate the collected information. This step was useful to (i) fill remaining gaps, (ii) interpret the obtained results, and (iii) interpolate and aggregate the large amount of collected information in a meaningful way. Expert knowledge also provides a quality assurance mechanism, enhancing the accuracy and reliability of evaluation findings but also enriching the evaluation with additional information and interpretation of data from desk research.

3. Results of the evaluation

This chapter reports the results of the evaluation for each sub-category of tools. The chapter is organised with a dedicated section for each category of tools, broken down in sub-sections for each corresponding sub-category of tools. For each category of tools, we provide a description of the tools, with examples from the inventory, and the about the policy tasks for which the tools were used. Following, for each sub-category of tools, we report a summary of results 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 is provided to bring 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.

3.1. Stakeholder needs assessment tools

3.1.1. Inventory of stakeholder needs assessment tools

Table 3 below shows examples of tools belonging to the different sub-categories, according to our inventory of tools. The full list of tools is available in the Tools4CAP online inventory.

Table 3. Sub-categories of tools and examples of tools from the Tools4CAP inventory.

Sub-category of tools	Example of tools identified
Online consultation and surveys	• E-savjetovanje (e-consultations) (Croatia) • Public consultation tool otakantaa.fi (Finland) • Online Survey (Croatia) • Online Consultation (Hungary) • Participatory Consultative Tool (Thematic Consultative Committee) (Romania) • Online Questionnaires (Romania)
Workshops and conferences	• Discussion Conference (Cyprus) • World Café (Finland, Germany) • Workshops, Conferences, Working Groups (Romania) • Consultation forum (Sweden) • Public Consultation (Belgium-Flanders) • Workshops and Round-Table Discussions (Hungary) • Network of Seminars (Latvia) • Workshop (Luxembourg)
Focus groups and meetings	• Town Hall meetings (Ireland) • Partnership meetings and Public Consultation (Spain) • Working Groups (Spain) • Stakeholder Consultation Rounds (Netherlands) • Internal Working Groups (Belgium-Flanders) • Internal Focus Groups (Cyprus) • External Focus Groups (Cyprus) • Open Public Consultations (Ireland) • Impactons! Public Debate (France) • Focus Group (Lithuania) • Interviews (Bulgaria) • Working Groups (Slovakia)

Stakeholder needs assessment tools have been employed to support different tasks of the CSP design and monitoring. Table 4 below provides an overview of the number of MSs where a specific sub-category of tools was employed for each task of the CSP design and monitoring. More details are available in the online inventory of tools[9].

[9] https://www.tools4cap.eu/tools/

Table 4. Use of stakeholder needs assessment tools across the CSP design and monitoring (no. of MSs in which the tools were employed for the specific task).

	Online consultation and surveys	Workshops and conferences	Focus groups and meetings
CSP design tasks			
Socio-economic context analysis	1	1	2
SWOT analysis	1	5	4
Needs assessment	1	3	6
Interventions setting	2	1	3
Target setting			
Financial allocations	1		
Ex-ante analysis and SEA			
Stakeholders' consultations	3	5	4
CSP monitoring tasks			
Performance review			4
Beneficiaries' compliance			
Evaluation			

The following section presents the results of the evaluation specific for each sub-category of tools.

3.1.2. Online consultation and surveys

Summary

***Accuracy**: limited to moderate due to sampling issues, reliance on stakeholders' perceptions (unstable and not objective), lack of traceability of input data, self-selection biases.*

Data accuracy in online consultations faces challenges due to sampling issues, reliance on subjective stakeholder perceptions, lack of traceability, and self-selection biases. This subjectivity introduces instability over time, compromising overall accuracy. Despite these challenges, online consultations capture diverse stakeholder perspectives, valuable in areas with limited prior knowledge. Stakeholders report employing rigorous measures such as data collection protocols and verification meetings to enhance accuracy. In terms of representativeness, online tools can engage thousands of stakeholders via ICTs, yet self-selection bias in public consultations may exclude certain stakeholder categories. Targeted consultations like public debates could mitigate this issue. Half of the survey respondents note that available tools facilitate the involvement of under-represented vulnerable stakeholder groups.

***Reliability**: moderate, these tools are easy to replicate (e.g. EU Survey), but it is difficult to get the same results due to sampling bias, self-selection, influence of size of the sample.*

The replicability of online consultation tools and surveys is feasible, particularly exemplified by tools like EUSurvey. However, it is crucial to acknowledge limitations, especially regarding sampling issues (sampling bias) that can hinder replicability like public consultations with self-selected respondents. Factors like unbalanced samples become more pronounced, affecting the results' replicability, particularly when participant numbers are low. Concerning transparency, objectives and results are usually public, however, information on how they serve the policymaking process is lacking. Final outcomes are often in processed forms, with less clarity on how results inform policymaking. The complex nature of consultation data challenges understanding its influence on policy processes, with mechanisms and influence remaining less transparent.

***Applicability**: high, they are easy to adapt to different contexts, policy steps and CAP SOs, there are few limitations e.g. in case of very complex topics (mistrust issues).*

Online consultation and survey tools are highly adaptable, capable of contributing to discussions on various policy objectives and steps in policymaking. Nevertheless, the intricacy of specific subjects or policy processes may limit their effective use in informing policymaking. Besides, in this CAP programming period, at least nine MS utilised online consultations for CSPs. The extent of past MS' use is less clear, but some, like Czechia and Slovenia, had adopted similar tools in the previous programming period. Additionally, online consultations are considered flexible as they can provide information at different granularity ranging from very broad public consultation to very stakeholder-targeted consultations.

***Accessibility**: high, they are easy to use, limited knowledge needed (e.g. survey design, technical skills).*

Online consultation surveys are generally user-friendly, requiring no extensive training, making them accessible to a wide range of professionals. While some instances required prior training, others found it unnecessary. Overall, the tools demonstrate versatility and ease of application, accommodating diverse users in the policymaking process. Online consultations are generally openly available, with various platforms and guidelines accessible to users. While the EC conducts numerous consultations, countries may use dedicated platforms or private survey services. Factors such as costs, complexity, and technical requirements can limit tool availability according to survey respondents. Additionally, only 20% of the survey respondents reported that documentation concerning the underlying methodology of the tool/documentation is available and accessible to all stakeholders.

***Efficiency**: high, they are quick to set up but often long implementation, cost effective (compared to individual consultations), relatively small teams needed for implementation.*

Setting up online consultation and survey tools is typically swift, taking a short duration, while implementation times vary but generally range from a few weeks up to six months. The efficiency of these tools in reaching a diverse audience in a short time is a notable advantage, particularly in comparison to traditional face-to-face methods.

Implementing online consultations and surveys is generally cost-effective, with 60% of the survey respondents estimating costs below 10,000 Euro. Compared to traditional consultation means, online tools are generally less expensive, particularly when utilising existing public e-platforms or government websites. Implementing online consultation and survey tools typically requires small teams, usually less than 5 people according to 40% of survey respondents. Automation aids in data processing, making these tools accessible to a wide range of professional backgrounds.

Analysis by criteria

ACCURACY

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

DATA ACCURACY AND VALIDITY

This sub-criterion refers to the extent to which data input and its sources are traceable and validated.

Evaluation Question: Are measures in place to avoid mistakes and redundancy and to ensure consistency in the stakeholder's responses or information provided (e.g. to verify the quality and reliability of the information)? Are results triangulated, corroborated, or substantiated by other data sources? Are different types of stakeholders involved?

One important aspect examined for assessing the precision of this sub-category of tools pertains to the accuracy and validity of the input data. Specifically, this sub-criterion focuses on the information utilised during the tool's implementation. In the context of online consultations and surveys, this input data comprises the responses and information provided by stakeholders to feed into the tools implementation (e.g. stakeholders' views, opinions, responses).

According to the results of the evaluation, the information and data collected through online consultations and survey tools can suffer from a number of limitations affecting their validity to properly inform analysis or policymaking. Yet, online consultation and surveys can be adequately complemented, or being a complement to, other participatory or analytical tools. Firstly, online consultations and surveys collects data based on stakeholders' perceptions, therefore not objective and variable over time (Bertolozzi-Caredio et al., 2023). Secondly, these tools are affected by sampling issues, notably respondents' self-selection and lack of knowledge about respondents' characteristics, which lead to unbalanced representativeness, no generalisation and low validity of data (Andrade, 2020[10]; Wright, 2005[11]). Besides, the validity of the collected data depends on the quality of the questionnaire and the respondents' understanding of it, and the completeness of the collected questionnaires.

Lastly, sources of collected information are not always traceable. In fact, according to the Tools4CAP survey results (Section 2.2.3 Survey), the source of information is traceable according to 60% of the respondents, while the remaining 40% declares that it would not be possible to trace back the information to its original source. According to 50% of the Tools4CAP survey respondents, MSs adopted specific measures to ensure the validity and accuracy of the information obtained through online consultations and surveys, whereas 30% declares that no measure was in place. These measures consist in *data collection protocols*, *automated data cleansing*, as well as *double-entry verification*, and *ad-hoc meetings with survey participants* to verify the validity of the information collected.

According to the focus groups results in France, the accuracy of the result is heterogeneous, as the outcomes of the consultation are mostly a result of individual contributions. In addition, Romania's focus group had a similar point of view, stating that multi-stakeholder online consultations may be often flawed due to the particular interests of some groups that attend the meeting and are not willing to identify a solution that can equally satisfy the needs of the other parties from the discussion table. According to our focus group, not all consultation tools had a validation step for

[10] Andrade, 2020. https://journals.sagepub.com/doi/10.1177/0253717620957496

[11] Wright, 2005. https://academic.oup.com/jcmc/article/10/3/JCMC1034/4614509

the results, as was the case for ImPACtons! in France. Meanwhile in Italy, some tools had a validation step through peer review, and some others did not.

Despite these limitations, online consultations, especially those based on open questionnaires, prove useful in gathering stakeholders' and citizens' views, opinions, and experiences on certain topics about which previous knowledge is lacking[12]. For instance, in Finland respondents were asked to comment on the different sections of the draft CSP[13], whereas in Croatia respondents were also asked to share their experiences and views about a specific topic, notably the use of the financial instruments[14]. Respondents were given access to material and documentation to get insights on the topic.

Still, online consultation and surveys can be profitably complemented with other tools, such as participatory approaches (e.g. workshops) to discuss and validate the consultation's results or quantitative modelling to corroborate them with empirical evidence. For example, in Romania, the online consultation was complemented by online thematic working groups to help structure the discussion around the needs assessment[15]. The German focus group indicated that having an online consultation that follows a stakeholders' conference can be very successful in capturing and reflecting stakeholder's views.

Tool Representativeness

This sub-criterion assess the capacity of Online consultation and survey to accurately and comprehensively capture or depict a system, dataset, or phenomenon. To have a better understanding of how representative online consultations and surveys are, the evaluation investigates elements such as the number of stakeholders involved in the implementation of such category of tools, the degree of participation of different stakeholders groups as well as potential biases affecting the tools.

Evaluation Question: How many stakeholders can be involved in the implementation of this sub-category of tools (e.g. number of participants of the workshop, focus group, consultation etc.)? Does this sub-category of tools facilitate the involvement of all relevant stakeholder groups while ensuring representativeness? Is this sub-category of tools affected by biases?

Online consultation and surveys can reach a relevant number of stakeholders (in the order of thousands at least), and potentially involve all categories of stakeholders (down to the single citizen). Yet, especially public consultations, might results in unbalanced samples in terms of geographical and stakeholders' categories representativeness. According to the Tools4CAP survey result, the vast majority of respondents (80%) reported that a number of at least 50 stakeholders have been involved, with consultations sometimes exceeding the 250 stakeholders (30%).

Online consultations and survey tools can potentially reach a large involvement of stakeholders from different levels of society. The number of consultations open to the whole public to participate has increased in OECD countries[16]. The use of ICTs, such as e-platforms and social media, has drastically increased the potential number of participants in consultations (Musial-Karg, 2019). Public consultations, in particular, can reach a very high number of responses. For example, in France, 9.506 contributions were collected through the e-platform Impactons[17], of which 7.409 were responses to an online questionnaire and the rest were other feedback such as commentaries on the website and social media. In Ireland, about 1,500 contributions were gathered through SurveyMonkey[18], whereas in Finland

[12] European Commission, 'Better regulation' toolbox 2023, available at: https://commission.europa.eu/system/files/2023-09/BRT-2023-Chapter%207-Stakeholder%20consultation_0.pdf

[13] Ministry of Justice of Finland, available at: https://www.otakantaa.fi/fi/

[14] https://savjetovanja.gov.hr/o-savjetovanjima/sto-su-to-e-savjetovanja-i-kako-se-ukljuciti-1123/1123

[15] Ministry of Agriculture and Rural development of Romania, available at: https://www.madr.ro/planul-national-strategic-pac-post-2020/programare-ps-pac-2023-2027/consultari-pns.html

[16] https://www.oecd.org/mena/governance/MENA-Practitioners-Guide-%20EN.pdf

[17] https://impactons.debatpublic.fr/wp-content/uploads/compte-rendu-d-etape-18062020.pdf

[18] https://www.gov.ie/en/collection/79383-cap-post-2020-public-submissions/

2.486 responses were collected through the e-platform otakantaa.fi[19]. These consultations were open to farmers, associations, NGOs, public administrations and, more generally, to all citizens.

However, the geographical and stakeholder representativeness of the results from online consultations and survey tools can be affected by self-selection bias[20]. This implies that there is no direct control over which individuals (or from which regions or categories of stakeholders) participate in the consultation, resulting in a potentially large sample not reflecting the reality. Public consultations are particularly affected by self-selection, whereas targeted consultations can mitigate or avoid this risk. Alternatively, public consultations can be complemented by other stakeholder engagement events. For example, public debates were organised in France and Ireland. There might always be some concerns surrounding the representativeness of the surveyed population as certain demographics may be underrepresented compared to others, due to the digital literacy of some participants, which can potentially affect accessibility[21].

While online consultations enhance the involvement of interested parties and formally offer the possibility to give qualitative input, they are often criticised for not being fully inclusive of less concerned categories of stakeholders (e.g. single citizens or non-purely agricultural NGOs), and transparent in the process of collection and analysis (Quittkat, 2011)[22]. However, according to the majority of Tools4CAP survey participants, this sub-category of tools can facilitate the involvement of vulnerable stakeholder groups which are usually under-represented (e.g. people with disabilities, ethnic minorities etc..) especially when the tool is implemented in an online environment.

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

This sub-criterion is used to evaluate 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 replicability of online consultation tools and surveys is in principles feasible due to the user friendliness, web-based nature and open-source availability of the common tools. However, certain sampling issues can hinder the replicability of consultations' results.

An example is EUSurvey, which is available in 23 official EU languages, making it suitable for Europe-wide online consultations and surveys. Moreover, its free public availability to all European citizens and the fact that it is easily and highly customizable makes it flexible and replicable. These mentioned characteristics make all steps of the survey process smoother, from design to analysis[23]. In addition, research has shown that web-based survey research is becoming increasingly popular and better accepted by researchers, indicating the high potential for replicability (2).

Nevertheless, it is important to note that there are some limitations in terms of replicability for online consultations and surveys. Additionally, it is important to note that sampling limitations can hinder obtaining the same results when

[19] https://www.otakantaa.fi/fi/hankkeet/331/

[20] Better Regulation Guidelines, available at: https://commission.europa.eu/system/files/2021-11/swd2021_305_en.pdf

[21] Olivieri S. et al., 2021, "*Opportunities and Challenges of Web-Based and Remotely Administered Surveys for Patient Preference Studies in a Vulnerable Population*", available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8613941/

[22] Christine Quittkat, 2011, "The European Commission's Online Consultations: A Success Story?", available at: https://onlinelibrary.wiley.com/doi/10.1111/j.1468-5965.2010.02147.x.

[23] https://ec.europa.eu/isa2/actions/tool-europe-wide-online-consultations-and-surveys_en/

replicating the consultation. For example, public consultations where respondents are self-selected (they choose whether to take part or not) can make the sample non-representative and thus researchers cannot foresee a guaranteed replicability due to potential unbalanced samples[24]. Self-selection can affect results especially when the number of participants is low, as higher participation might lead to sufficient randomisation of the sample.

According to a Polish focus group, the consistency of findings can depend on the reliability of stakeholders and their readiness to present consistent opinions irrespective of the political context. While a German focus group stated that online consultation and survey tools offer consistent findings but are also very context-specific. They indicated that a success factor for ensuring long-term involvement and preventing a tiring of stakeholders is seen in follow-up conferences or workshops that are built on the survey results.

Transparency

This sub-criterion is used to assess 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?

Objectives, sampling and results of online consultations and surveys are generally publicly available and transparent, although this can be done in different forms, while it is less clear how consultations' results have been brought into and informed the policymaking process.

Online consultation and survey tools normally do not rely much on complex methodologies or protocols, whereby there is usually no need to present extensive technical or methodological protocols for replication. It is however important in online consultation to be transparent regarding the objective, the sampling strategy and target population, the means for data collection, and the analytical or data processing methods, as well as the way results are brought into the policymaking process (4). Furthermore, it is necessary both to account for which comments contributed to the final decision and to give reasons for the acceptance and rejection of the positions expressed by the participants (5). It is critical for respondents to know how and to what extent their input is taken into account and it is especially important to explain why certain suggestions could not be taken up in the policy formulation (6).

Most often, objectives, samples' characteristics and results are publicly available, although final results are often published in a processed or aggregated form, like in a report. For instance, France provides a report with mostly processed data (1), whereas in Ireland the collected questionnaires are publicly available (2), and in Finland the collected comments to the draft CSP were public (3). Therefore, there are different ways the results can be made transparent. Nonetheless, only 20% of the respondents to our survey reported that stakeholders have certainly access to the documentation concerning the tools' underlying methodology.

There is, however, less information available regarding how the results of the online consultations informed the design of the CSPs. Because of the complex nature of consultations' data and results (often gathered in mixed forms, with massive amount of qualitative information hard to process), it is more difficult to define the way, extent, and moment in which this information enter into play in the policy process, which in turn makes harder to clarify (i.e. make transparent) the actual role of this information. Most often, it is explained that policy outputs have been built based on consultations' results, but it is not clear how and to what extent these results have actually informed the policy output.

[24] https://commission.europa.eu/system/files/2023-09/BRT-2023-Chapter%207-Stakeholder%20consultation_0.pdf

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

This sub-criterion is used to measure 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?

Overall, online consultation and survey tools are very flexible and can be used to inform the discussion around any policy objective and policymaking step. However, the complexity of certain topics or policy processes might hinder the actual capacity of online consultation to properly inform policymaking, whereas the public may also be reluctant towards the use of public consultations due to mistrust (1).

For example, according to the online survey we conducted, online consultation tools were used to address CAP objectives related to SO1 (fair income), SO7 (generational renewal), and CCO (knowledge and innovation) in Czechia, SO4 (climate change) and SO6 (biodiversity) in Slovenia, and SO1, SO3 (food value chain), SO6 and SO9 (food and health) in Poland. Potentially, online consultation tools could be adapted to respond to any objective.

Similarly, according to the respondents to our online survey, online consultation tools could be used to inform any of the CSP design steps. In Czechia, for example, these tools were used to inform the ex-ante analysis, the needs assessment, and the design of interventions and targets, while in Poland they were also used to inform the socio-economic context analysis and the budget allocation.

According to our focus groups results, online consultations and surveys are mostly flexible. For the German focus group, stakeholder consultation tools can be adapted to all contexts that rely on the assessment of stakeholder's needs. This includes all policy steps in which the involvement of stakeholder's needs is foreseen and all specific objectives. While for the 3 tools that had focus groups in Italy, it was agreed that they could be used for other steps such as needs prioritisation, needs assessment, SWOT analysis, and ex-ante evaluation.

Experience

This sub-criterion is used to measure the extent to which the tool has been used across MS, in previous programming periods, or in other policies.

Evaluation Question: How experienced are MS 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?

Online consultation and survey are well-established tools in policymaking, in particular widely used by the EC not only for agriculture policy (e.g. Ammann and Boussat, 2022), which also provides several examples and guidelines. However, it is less clear how MS made use of these tools in the past.

According to the interviews and our online survey, at least nine MSs have made use of online consultations to support the design of the CSPs in this programming period, including, France, Ireland, Hungary, Romania, Finland, Croatia, Czechia, Slovenia and Poland. Other MS have adopted other participatory approaches. To note that, in this programming period, stakeholder engagement was required by Regulation 2115/2021 (1).

For the past programming period (CAP 2014-2022), the EC did conduct consultation activities open to stakeholders and the wider society (2, 3). However, it is less clear the extent to which MS conducted online consultations. According to the respondents to our survey, the tools adopted in Czechia and Slovenia were also adopted in the previous programming period. In other cases, MS might have relied more on the engagement of specific advisory and stakeholder groups at national level, whereas the more open public consultation were conducted by the EC. Yet, MS could have conducted online consultations on specific topics. For example, Italy conducted a telematic open consultation (via email) on the Operational Programmes in 2014. The novelty of Reg. 2115/2021 requiring stronger stakeholder engagement might have led to a wider use by MSs of online consultations compared to the past.

Granularity

This sub-criterion is used to measure 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? What is the level of detail that these tools can reach?

Online consultations and survey tools can be adapted to very different geographical scopes, and range from very broad public consultation to very stakeholder-targeted consultations. Overall, these tools are flexible and can provide information at different granularity. If the relevant stakeholders are few (i.e. not the general population), one may need to move from a logic of large samples to ensure representativeness to a logic of careful selection of all relevant stakeholders at lower geographical scales. For example, Czechia and Poland conducted more targeted consultations that involved less than 250 stakeholders (according to the respondents of our survey), whereas in Ireland and France about 1,500 and 7,400 responses were gathered respectively.

When asked about the governance levels involved in online consultations and surveys, the majority of respondents to our online survey (60%) indicated that these tools have been used to engage with stakeholders at different governance levels. According to our focus groups results, different tools had different levels of granularity. For example, France's ImPACtons! can be used at the local, regional, and national levels. On the other hand, in Italy, the used online questionnaire could only be used either at local or regional level.

The granularity of the tool depends on the type of approach used. In the first place, consultations can be geographically binding. However, more open approaches tending to online consultations are typically less targeted, whereas more closed approaches (e.g. workshops and expert groups) are usually more targeted and used to get higher granularity on specific topics or sectors. Hybrid approaches, combining open and closed solutions, can take the benefits of the two ensuring both wider societal coverage and high-granularity information (Fraussen et al., 2020).

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

This sub-criterion aims at assessing 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?

Online consultation and surveys are in general straightforward and easy to use, they do not need any prior intensive training or experience. This leads to a bigger availability of the tool to a wider range of stakeholders and citizens. Three out of four respondents to our survey (from Czechia and Slovenia) felt very satisfied with the ease of use of these tools, while one respondent from Poland felt not satisfied about it. In Czechia and Poland, using the tool required priori training (between 20-50 hours) whereas this was not needed it Slovenia. According to our focus group results, online consultations and surveys were generally considered easy to use. Germany's focus group added that these tools demand competencies in survey question design, the technical tools do so as well as capacities to recruit

participants, analyse the data and follow-up regarding the specific aim of the survey. Furthermore, France's focus group stated that the implementation of the tool is easy, however taking the important number of propositions into account was an arduous task for the managing authorities.

The following are examples highlighting the ease of use of these tools. In the case of Finland, the Ministry of Justice has an official online website (otakantaa.fi)[25] that has been in place for a while. It consists of a tool for collecting public consultations about projects from different organisations. For the CAP, the FI SP draft was divided into parts and full texts were provided in PDF format. Stakeholders could easily access, read, and comment on different sections and ask questions. The comments are visible to the public, encouraging open discussion and transparency. Additionally, Croatia has a similar process where documents that are under consultations are published on a specific website, available for commenting, and after 30 days ministries and agencies are obliged to publish a report with all comments and justifications (e.g. accepted or not, and why) (3). Stakeholders can also subscribe to certain topics in order to be alerted once newly opened consultations are published. In both cases, the ministries and the users are now used to this kind of consultation, since they have been doing it for a while.

In Romania, Technical Working Groups (TWG) participants were provided with an online questionnaire asking them to choose between different options for the CSPs interventions and to express their views on interventions and budget allocation. The Romanian Ministry of Agriculture (MADR) chose this online questionnaire tool for its ease of application and the ability to quickly analyse results[26]. The computerised database with answers allowed for efficient iteration making it easier for end users to process the collected information. Additionally, MADR, through the AKIS coordination unit organised consultations with all representative actors and the Support Unit to approve the topics of interest (MADR, 2022).

Tool Availability

This sub-criterion assesses 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?

Several platforms and guidelines are available for the implementation of online consultation and survey tools. Online consultation and survey are well-know and long-standing tools in science and public consultations. Several technical and methodological guidelines are available based on examples from scientific research or policy consultations. For example, the Better Regulation Toolbox of the EC provides methodological indications for implementing consultations and surveys. Relevant guidelines are also provided by the United Nations[27]. The EC itself conduces numerous consultations[28].

However, consultations can be conducted through different means, namely dedicated e-platforms, web sites or private survey platforms. For example, Ireland made use of SurveyMonkey, a private survey service, while France set up a dedicated platform called "Impactons!". Private survey services are available at a price. There are several existing platforms on which end users can rely, including the freely available EUSurvey platform set up by the EC. Setting up a dedicated platform can be costly and technically complex but in principle always feasible.

Normally, there are no particular legal or patent-related limitations constraining the availability of online consultation and survey tools. Costs, complexity and technical requirements can be important limitations according to four respondents to our online survey.

---

[25] https://www.otakantaa.fi/fi/

[26] https://www.madr.ro/planul-national-strategic-pac-post-2020/programare-ps-pac-2023-2027/consultari-pns.html

[27] https://www.undp.org/sites/g/files/zskgke326/files/migration/vn/31473_Manual_on_public_consultations_for_PPCs-eng.pdf

[28] https://commission.europa.eu/about-european-commission/service-standards-and-principles/transparency/consultations_en

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

This sub-criterion measure 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 sub-category of tools (i.e time necessary to organise and prepare the implementation, including the identification and invitation of participants, setting up the online page of the consultation or - in case of methods implemented in presence - the arrangement of the venue, etc.)? And how much time does it take to implement the process of the tool (i.e. time necessary to implement the tool until the achievement of the intended outputs)?

Setting up online consultation and survey tools is a relatively swift task. The duration of the implementation can vary depending on the targets and the performance of the campaign, but remains anyway limited within few months.

According to the respondents to our survey, setting up these tools required one month in Poland, between 1-2 weeks in Czechia, and less than one week in Slovenia. This seems to indicate high time efficiency in terms of the tool's set-up. Regarding the implementation, it varies from a few weeks in Slovenia and Czechia, to one to six months in Czechia (a different consultation) and Poland.

The focus group results reflect different opinions about the time efficiency of online consultations and surveys. Romania's focus group stated that the time allocated to the public consultation process and synthesis of feedback received from stakeholders was quite significant, as both the CSP design process and the public consultations went in parallel and were managed by the CSP Managing Authority in the Ministry of Agriculture. Similarly, Italy's focus groups indicated that the implementation time was long. On the other hand, Poland's focus group indicated that consultations and online survey were considered to be the least time consuming.

However, online surveys, e.g. used in research, allow the researcher to reach thousands of people with common characteristics in a short time, despite possibly being separated by great geographic distances (Bachmann & Elfrink, 1996; Garton et al., 2003; Taylor, 2000; Yun & Trumbo, 2000). In the face-to-face research environment, it would take considerably longer-if it were possible at all-to find an equivalent number of people with specific attributes, interests, and attitudes in one location. Online surveys may also save time by allowing researchers to collect data while they work on other tasks (Llieva, Baron, & Healey, 2002) and due to the possibility of immediately transmitting responses via e-mail, it allows preliminary analyses to be conducted on the collected data while waiting for the desired number of responses to accumulate (Llieva et al., 2002) (3).

Cost-effectiveness

This sub-criterion measures the extent to which the tool is able to achieve its outcomes within a reasonable amount of 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 budget can be necessary for the maintenance of the tool or subscription for patent-based tools every year?

Online consultations and surveys can be very cost-effective tools to reach a wide target of stakeholders. The four respondents to our survey that described online consultation tools, when asked to what extent costs can hinder the access to the tool from 1 to 5 (5 being the highest constraint), scored all 3 or below. In all cases, the estimated cost was lower than 10,000 Euro, although a respondent from Poland specified that the subscription to a private tool would have costed cost between 10,000 - 50,000 Euro annually.

According to our focus groups results, online consultation and surveys were generally cheap. Romania's focus group indicated that the costs consisted of only the ''Zoom'' subscription, while Poland's focus group stated that surveys

proved to be the cheapest because the Ministry's existing website was used. Furthermore, the Italian focus group declared a total cost of less than 1,000 Euro for each of their online questionnaires. France's ImPACtons! was the only tool that was declared budget consuming (as well as people- and time- consuming).

A more detailed budget report was published in France (2), which however included not only the online consultations, but also several public debates and social media campaigns. The budget reached 1,356 Euro spent, of which about 13% for managing the e-platform, and the rest for logistics, personnel, and communication.

Compared to other consultation means (e.g. face-to-face, on spot, etc.), and to reach the same target stakeholders, online consultation and survey tools are generally less expensive (Wright, 2005)[29]. Expenses like travel, materials, transcriptions, or telephone are avoided. Higher costs might emerge if dedicated e-platforms have to be set up. However, several private platforms are available at accessible costs, whereas some countries (like Finland) relies on already established public e-platforms. Often, the online consultations have been launched on government websites and gathered through institutional emails (e.g. respondents have to send comments by email), which is even more cost-effective, although this can decrease response rate and/or increase complexity in managing the gathered data.

Human Resources Efficiency

This sub-criterion measures the extent to which the tool is able to achieve its outcomes involving a reasonable amount 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?

Online consultation and survey tools can be run by relatively small teams, and generally there is no need of highly specialised professionals. According to the respondents to our survey, the implementation of the tool required less than 10 people in Poland, and less than five in Slovenia and Czechia. When prior training was required, this was limited to 20-50 hours in Poland and Czechia, and between 1-20 hours in Lithuania. In France, a team of eight people was required (2). One of the survey's respondents, though, flagged that compared to tools implemented in presence, online consultations require availability of ad-hoc technical support. Therefore, the knowledge of the online platforms used to host such consultations is required.

Setting up an online consultation requires the definition of a stakeholder target and a questionnaire, as well as the planning of collection and process monitoring, and data processing. Once the consultation has started, it is necessary that the team dedicated to it closely follows the development, both supporting the promotion activities, and tracing provisional statistics on participation and the content of the comments received. Especially in the case of online consultations, it is necessary to check, from a technical point of view, the correct course of the initiative, in order to correct any unforeseen or errors. It is good practice to periodically estimate the extent of participation and to proceed with the analysis of the textual contents of the contributions received, in order to understand in good time the meaning of the information coming from the participants (1). In principles, these activities can be conducted by small teams, especially thanks to the automatization of many data processing tasks. Bigger and more specialised teams might be required if a dedicated e-platform has to be set up.

[29] Wright, 2005. https://academic.oup.com/jcmc/article/10/3/JCMC1034/4614509

3.1.3. Workshops and conferences

Summary

***Accuracy**: limited to moderate, due to sampling issues, representativeness affected by self-selection in some cases and openness bias, measures to ensure quality input data quality not implemented systematically.*

Workshops and conferences may encounter limitations in data quality and validity, especially if participant diversity is not reflective of target groups. The outcomes' quality depends on organiser settings, participant selection, and facilitator expertise. Online formats might yield fewer open responses, while hybrid formats can lead to imbalances. Additionally, 50% of the survey's respondents indicate that there are no measures in place to ensure accuracy of the information provided. Despite these challenges, the two tools can still complement other tools, contributing to a comprehensive understanding when effectively organised and utilised. Workshops and conferences, with varying participant sizes, involve selective invitations or open invitations, impacting representativeness. Open invitations may lead to non-representative participation. The challenge lies in balancing inclusivity and selectiveness in participant selection. Smaller groups may lack representativeness, while larger workshops require suitable methods to avoid imbalance.

***Reliability**: limited, due to lower replicability than surveys, established methodologies are not implemented systematically and not accessible, openness of participants affected when results are made public.*

Replicability of workshops and conferences relies on participant representativeness and the chosen assessment methods. Due to deliberation and political factors, replicability is less assured compared to surveys. The openness of the chosen approach influences the likelihood of successful result replication. Only 57% of survey respondents indicated that workshops and conferences are based on established methodologies and known steps. The transparency of this tool's underlying methodologies, data, and results depends on the availability of documentation. Full transparency can be achieved through methods like transcription and videotaping, potentially hindered by participants' reluctance in public expression. Additionally, the quality of outcomes depends on facilitator and team capacity, with the communication of results to participants being a standard practice, though public availability might vary. 86% of survey respondents either declare that no detailed documentation for this subcategory of tools is available and accessible to all stakeholders (54%) or are not aware of such documentation (31%).

***Applicability**: high, as they are easy to adapt to different contexts (including different governance levels), policy steps (especially during design) and CAP SOs, widely used by MS in previous programming periods.*

Workshops and conferences, as consultative participation tools, are versatile for addressing a wide range of policy objectives and policymaking steps. They have been employed to identify stakeholder needs and work towards creative solutions in various stages of the policy cycle. Workshops are more widely used than conferences. The vast majority of survey respondents (82%) used this subcategory of tools in the previous programming period. Workshops and conferences can be customised to address specific questions and methods, making them adaptable to support analysis and policymaking at different granularity levels, from local to regional, national, and EU levels. The majority of survey respondents (72%) consider that workshops and conference tools support the engagement of stakeholders at various governance levels (mostly national and/or regional level).

***Accessibility**: high, they are easy to use, limited knowledge needed (e.g. facilitators, experience of participants), widely and openly available*

Facilitating workshops and conferences requires skilled individuals who can maintain a neutral position, handle conflicts, and ensure diverse participation. However, the larger the group, the more complex the topic, and the more controversial the opinions, the greater the need for experienced facilitators to ensure effectiveness. More than half of survey respondents (around 60%) consider that workshops and conferences do not require any prior training or experience. Workshops and conferences tools and approaches as well as trained facilitators are widely available. However, some factors such as experience of facilitators, costs for external facilitators and organisers, and willingness of relevant stakeholders might limit the tool's accessibility.

***Efficiency**: moderate, depending on the size and methodology, higher costs, more time to prepare, and more human resources when the event is physical compared to online tools.*

Overall, the efficiency of workshops and conferences depend on the methodological design and the format. In terms of time efficiency, workshops typically require 6-8 weeks of preparation, while conferences aimed at larger audiences need a longer lead time, starting preparations at least 6 to three months before the event. Implementing workshops or conferences is generally cost-effective, mainly requiring budget for personnel, facilitators and, in case of physical events, venue, equipment, consumables, catering, and travel costs, without significant expenses for patents or licenses. Workshops and conferences generally require at least one professional to define topics, an organiser for logistics, and an optional external facilitator, with additional personnel for documentation. IT support might be required for virtual or hybrid settings. The majority of survey respondents declared that less than 10 people were necessary to implement the tool.

Analysis by criteria

ACCURACY

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

DATA ACCURACY AND VALIDITY

This sub-criterion refers to the extent to which data input and its sources are traceable and validated.

Evaluation Question: Are measures in place to avoid mistakes and redundancy and to ensure consistency in the stakeholder's responses or information provided (e.g. to verify the quality and reliability of the information)? Are results triangulated, corroborated, or substantiated by other data sources? Are different types of stakeholders involved?

The information and data collected through workshops and conferences can suffer from a number of limitations affecting their validity to properly inform analysis or policymaking. If, for example, the participants in a workshop do not reflect the actual diversity of the target groups, or the formats used do not sufficiently compensate for imbalances, the results of the workshops may be unbalanced. Yet, workshops and conferences can be adequately complemented, or being a complement to, other participatory or analytical tools.

The quality of the outcomes depends largely on the chosen settings and methods of the organisers, the choice of participants, and the expertise of the facilitator(s). The format can also influence the results – online formats (as opposed to physical formats) might lead to less open responses of participants as it is more difficult to build trust among participants. Hybrid formats (some participants are online, some physical) can lead to unbalanced results .

Unlike a non-verbal interaction, an oral hearing inherently suffers from the lack of accuracy. Despite the fact that workshops are getting increasingly IT-oriented, agents partaking in a workshop have a limited access to hard-wired data. While presentations are on the go, speakers might be tempted to skip contested points, which leads to a dwindling interest on the part of the audience. Both events lead to a loss in information content between the encoder and the decoder side according to Shannon's concept of information entropy, ultimately leading to the increase in the chance of information loss or miscommunication (Vogels et al., 2020).

Besides, the result of the survey suggests that measures to avoid mistakes and redundancy in the stakeholder's responses or information provided (e.g. to verify the quality and reliability of the information) are not implemented systematically. In fact around half of the respondents indicates that such measures are not implemented. When implemented these measures usually consist of standardised data collection protocols or double-entry verification. One of the respondents declared that the outputs of the workshops go through a qualitative analysis and the review of the recordings, in order to consolidate and verify the correct understanding of the results emerged during the workshop implementation.

According to our focus groups results, different groups have different perspectives about the accuracy of workshops and conferences. Poland's focus group indicated that accuracy depends on how the different tools are implemented,

although workshops and conferences/seminars are the least accurate due to the possibility of being dominated by one person. On the other hand, the France's focus group indicated the results are accurate.

Tool Representativeness

This criterion refers to 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: How many stakeholders can be involved in the implementation of this sub-category of tools (e.g. number of participants of the workshop, focus group, consultation etc.)? Does this sub-category of tools facilitate the involvement of all relevant stakeholder groups while ensuring representativeness? Is this sub-category of tools affected by biases?

While workshops are typically only open to invited participants, conferences are usually more open. Participants may then self-select to a higher degree. However, there are also setting where participants are selectively invited. The question of selection and openness in terms of participants has consequences for representativeness.

The number of participants of workshops and conferences varies across a wide range, depending on the chosen settings. Workshops are most commonly organised with 10 to 25 participants; conferences can be much larger, and in this context ranging from 50 to something like 300 participants. Based on our survey, workshops have been implemented for the design of the CSPs involving a limited number of participants (less than 25 participants) in most cases. However, a relevant number of respondents report on events involving more than 50 participants, especially in situations where stakeholders from multiple governance levels (i.e. national, regional, sub-regional/local level) are invited to participate together. According to Germany's focus groups, the Regional conferences involved small groups of stakeholders (5-7 stakeholder groups), while the stakeholder conference reached 110 participants.

Open invitations (as opposed to selected "closed" range of participants) are more likely to result in non-representative participation, thus results are not as likely to be repeatable. The size of the participant group furthermore influences the outcomes. The smaller the group, the less representative it can be for a larger topic. Opposed to that, results of larger workshops could also be unbalanced if the choice of methods used and the capacity of facilitators is not suitable for such setup.

Specific stakeholders may be thought of as representatives of certain groups, and particularly the selection of a limited number of participants for a workshop requires selectiveness that will usually be guided by a notion of representation. However how well participants speak for the interests of a group they are thought to represent is often questionable (Jastram & Berberyan 2023). This difficulty can stand in the way of the goal of collecting all relevant needs. Careful selection of stakeholders to involve may mitigate this pitfall. Workshops represent a "closed approach" in that stakeholders are pre-selected (Fraussen et al. 2020, 477), and are thus only as inclusive or comprehensive as the pre-selection. There is an incentive to invite only stakeholders that can be expected to bring little controversy or divergent needs into the process (Fraussen et al. 2020, 477). On the other hand, the pre-selection of participants in such formats can be and often is used to prevent undue dominance of, for instance, business interests, which can manifest in more open formats, where participants self-select (Fraussen et al. 2020, 488). There might be certain stakeholder groups that are less willing or capable to join these formats (e.g., due to time constraints, language issues, lack of interests in the policy process), they might be harder to reach and include in a representative way.

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?

Replicability depends largely on the representativeness of participants as well as on the chosen methods for assessment. In that there can be a degree of deliberation, consensus-building, and other political factors at play, replicability is likely less than in surveys – results are to some degree contingent on the dynamic that unfolds between the participants.

There are a broad range of methods and approaches that can be used in workshops to analyse stakeholder needs, providing different levels of participatory engagement in the analysis. These range from methods that just ask for confirmation of predefined needs to fully bottom-up of identified needs taking into consideration different stakeholder groups perspectives. The less open the chosen approach, the more likely the results can be replicated.

Besides, according to most respondents to our survey (around 57%), workshops and conferences are usually based on established methodologies and protocols, which helps replicability. Around 24% of the respondents disagree, while the remaining 19% is not sure or does not know.

In our focus group, different opinions were reported. For Germany's focus group, focus group and meeting as well as workshop and conference tools produce consisted findings for the specific moment in the policy-design process, yet cannot offer consistent findings understood as identical findings over time. Participants highlighted the relevance of the exact planning and structure of how the tools are implemented. Facilitation, open discussion formats and an early communication of meeting agendas to enable preparation for stakeholders are seen as success factors to ensure that there is a consistent input from stakeholders. On the other hand, in Poland's focus group, the lowest score for reliability was given to workshops and conferences reaching only 2.8. The participants stated that consistency of findings can depend on the reliability of stakeholders and their readiness to present consistent opinions irrespective of the political context.

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?

Transparency of the process and contents brought to light depends on documentation. Transcription, videotaping etc. provide full transparency in that regard. Even more transparency can be achieved by opening the events to journalists. These measures may, however, stand in the way of openness from the participants, i.e. as participants would not want to express their opinion openly fearing that this would be made public.

Transparency of selection criteria and selection process of stakeholders to invite are crucial (Keating, 2003). Due to the importance of the selection for representativeness (see above), the procedural legitimacy of the tools rests on this. The methodology is transparent in that workshops and conferences are a method widely used in many different contexts and purposes[30]. A range of guidelines are available. Documentation of the outcomes is an important part of the tool. However, the quality of the outcome depends largely on the capacity of facilitator and organising team. Generally, the methods require outcomes to be communicated to all participants. However, in some circumstances, the outcomes might not be made publicly available.

According to our survey, however, when it comes to workshops held for the design of the CSPs the majority of respondents (86%) either declares that no detailed documentation for this subcategory of tools is available and accessible to all stakeholders (54%) or is not aware of such documentation (31%).

[30] https://theworldcafe.com/key-concepts-resources/world-cafe-method/

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?

Workshops are considered a tool of consultative participation (Pisano et al., 2015). Both workshops and conferences can be used to identify stakeholder needs related to a very broad range of policy objectives and in many steps of the process. According to our survey, workshops and conferences seem to have contributed to address topics related to all CAP specific objectives during the preparation of the CSPs. Also, this specific sub-category of tools has mostly been used for need assessment, SWOT analysis, interventions setting, and socio-economic context analysis. Respondents suggests that workshops and conferences could be useful for more policy tasks, for instance in relation to financial allocation and targets setting.

In addition to stakeholder needs assessment, workshops have been at times used to work with stakeholders further along the policy cycle, such as in Germany, where one further goal of workshops in the World Café format was working towards finding creative solutions[31], and in Hungary, where the rationale for the use of workshops and round-table discussions included consensus-building[32].

The choice of these methods is often a question of time and resources. For example, conferences need a longer preparation period and substantial financial resources, while workshops can be organised usually with much less time and less resources needed, even online.

Important to note: these types of engagement often raise expectations at the side of participants, and the request for a response to identified issues and needs. If this is not taken seriously, stakeholders might feel frustrated or might lose interest for future collaborations.

EXPERIENCE

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

Evaluation Question: How experienced are MS 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?

Workshops have been widely used in MSs, also in other policy areas. Conferences are not used as frequently to collect stakeholder needs. One exception are the German federal states of Bremen and Lower Saxony, where in early 2021, an online conference with 118 participants was held on the EAFRD concept 2023-2027 (ML, 2021).

According to our interviews and survey, at least 10 MSs have used this sub-category of tools during this programming period, including: Cyprus, Czechia, Hungary, Lithuania, Poland, Portugal, Romania, Slovak Republic, Slovenia,

[31] https://www.tools4cap.eu/tool/world-cafe-2/

[32] https://www.tools4cap.eu/tool/workshops-and-round-table-discussions-in-person-and-online/

Spain. Besides, this sub-category of tools was used in the previous programming period in at least three MSs: Poland, Czechia and Spain.

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? What is the level of detail that these tools can reach?

As representativeness is not approached through the number of participants but the choice, it is also feasible to have suitable participant selection both at very small scale or at very large levels (EU). The questions addressed and method used in workshops and conferences can be tailored to the very specific requirements, thus it can support analysis at very different scales. In fact, according to our survey, the majority of respondents (72%) suggests that workshops and conference allows the engagement of stakeholders at various governance levels (mostly national and/or regional level). For instance, the network of seminars in Latvia was held at a local scale (Ministry of Agriculture of the Republic of Latvia, 2023), whereas the Swedish consultation forum took place at a national scale[33].

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?

Depending on the purpose of a workshop, there is more or less preparation and experience required. Generally, it is recommended to involve a facilitation expert, best external, to ensure an objective approach is implemented in a workshop. More than half of the survey respondents (around 60%) consider that this sub-category of tools does not require any prior training or experience. Yet, some respondents suggest that up to 20 hours of training can be necessary to be able to implement the tools properly. Depending on purpose and content of a workshop or conference, these formats can also be demanding for the participants, in terms of preparation, expertise and networks (Fraussen 2023, 155).

France's focus group indicated that the implementation of workshops and conferences is generally easy. However, a difficulty encountered was the allocation of sufficient time to stakeholders to prepare the meetings and to make suggestions to the propositions of the Managing Authorities. On the other hand, Poland's focus group gave the lowest score for accessibility to workshops and conferences (3 out of 5).

Workshops that focus on a narrow subject, and that involve participants that know each other well can also be implemented with less experience and without professional facilitator and bring good results. The larger the groups, the more complex the topic and the more controversial the participants opinions are, the more experienced the facilitator should be. Skilled facilitators take a neutral position and enable all participants to voice their opinions, and also to handle conflicts. If facilitators are chosen from inside Managing Authorities, objectivity can be hampered. Also, a lack of experience / capacity can lead to unbalanced results.

TOOL AVAILABILITY

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

[33] https://www.opengovpartnership.org/members/sweden/commitments/SE0016/

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?

Tools and their methodologies, as well as trained facilitators, are widely and openly available[34]. The choice of suitable approaches and formats for conferences and workshops however are also depending on contexts, as there are quite different cultures of participation in the EU (Keating, 2003). It is therefore helpful to involve facilitators who are familiar both with general culture in the respective region as well as with the CAP specific context of target groups.

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 sub-category of tools (i.e time necessary to organise and prepare the implementation, including the identification and invitation of participants, setting up the online page of the consultation or - in case of methods implemented in presence - the arrangement of the venue, etc…)? And how much time does it take to implement the process of the tool (i.e. time necessary to implement the tool until the achievement of the intended outputs)?

Workshops potentially can have shorter preparation periods than conferences, but always depending on the chosen format (i.e., physical, hybrid or online), length of the event (mostly half or full days if physical, two hours to half day if online) and number and availability of desired participants. When asking about efficiency, it should also be noted that the pure working time required to prepare a workshop or conference is considerably less than the time required ahead of the event to organise the room, facilitation and, if necessary, the selection and invitation of participants.

According to our survey, all respondents but one consider that less than 2 weeks are sufficient to organise a workshop (the one suggests that at least one month would be necessary). When it comes to the implementation time of the tool, 36% of respondents declare that these tools are implemented in between a day and a week time, but a couple of respondents suggest that the implementation can require one to six months.

According to the focus group results in France and Poland, the tool can be more time consuming. France's focus group stated that the amount of time needed to organise and prepare the meetings, to compile outcomes, to draft conclusions and to reach a compromise is significant. While Poland's focus group indicated that that the most input demanding tools were workshops and conferences. Such events must be carefully planned in advance and the invitations to participate in such events must be sent early enough to ensure participation of the invited stakeholders. On the other hand, Germany's focus group indicate that this sub-category of tools needs little to medium preparation time (up to 3 months).

The following elaborations can only be understood as generalisation, but times can differ largely. In the case of workshops, the invitation or safe-the date information to participants should be circulated at least 6-8 weeks before the event. Conceptualisation of the agenda and selection of participants should take place before that. The time for organisation depends on setup and number of participants. At least 3 working days should be calculated for agenda preparation, invitation, etc., some hours to 2 days for documentation and analysis. Conferences aim at a larger number of participants. The target groups of conferences are usually defined, but participants are not preselected; instead invitations are spread more openly. Due to the larger size, conferences need more time ahead of the event

[34] https://participatory.tools/tool-kit/, https://www.sessionlab.com/library/

to organise locations and to advertise the event. At least 6 to three months before date seems a reasonable timeframe to start the preparation process of a conference.

Cost-effectiveness

Cost-effectiveness refers to the extent to which the tool is able to achieve its outcomes within a reasonable amount of 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 budget can be necessary for the maintenance of the tool or subscription for patent-based tools every year?

Costs occur in relation to personnel (preparation, implementation, documentation and analysis), external facilitators (e.g., as subcontractors), meeting room(s), equipment (rental of technical equipment such as projector, online conference technology, pinboards) consumables such as paper, pens etc.; catering for participants, travel costs for participants. According to France's focus group, the workshops involving National and regional managing authorities required little resources, while the National cross-sectional dialogue with all stakeholders was budget consuming. In the case of Poland, our focus group indicated that the workshops were the most labour-intensive because they often required a tender procedure.

Some articles analyse the overall costs (and benefits) of stakeholder participation. They consider the entire process for all who are involved and provide a more general view on efficiency of stakeholder engagement (Anggraeni et al., 2019). According to our survey respondents, the cost of implementation of these tools is generally under 1,000 Euro, with few exceptions of workshops costed between 1,000- 5,000 Euro or more than 10,000 Euro. Only one survey respondents reported that maintenance or a patent was needed. However, the costs associated to such maintenance or patent was less than 1,000 Euro.

Human Resources Efficiency

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

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

According to our survey, less than five persons were needed to implement workshops and conferences for the preparation of the CSPs. Besides, the respondents indicate that no particular training or experience is needed (except on respondent, who suggests that up to 20 hours training might be beneficial). However, our focus group in France flagged that stakeholder conferences can be demanding in terms of required people.

The literature suggests that the amount of people necessary for the implementation of the tool and the analysis of results vary according to the number of participants, number of focus groups and meetings, the length of the workshops (Gell, 2023). The main professionals to account for are the organiser, the facilitator, and the data analysts.

3.1.4. Focus groups and meetings

Summary

Accuracy: *moderate, focus groups and meetings allow traceability and validation of data since sometimes data collection protocol are in places, but face challenges in accuracy due to varying public record access and potential lobbing biases.*

MSs implement a variety of methods to gather and verify the information in focus groups and meetings. Data sources are often traceable and validated, as the content of discussions is typically recorded and transcribed into protocols, although there is variation in the detail and public accessibility of these records across countries. About 30% of respondents to the survey reports the use of standardised data collection protocols to reduce mistakes and redundancy. Organising thematic focus groups with diverse stakeholders is a central strategy in policy development, varying from public discussions to expert ministry groups, as seen in countries like Ireland, Belgium-Flanders and Cyprus. This approach promotes innovative perspectives but also poses a risk of bias from lobbying interests, highlighting the need for effective management to ensure objectivity.

***Reliability**: moderate, the tool's results are easily replicable but vary due to different factors, including national methodologies extending the EU guidelines provided in the methodological material for CSPs' preparation.*

Most survey respondents report that focus groups and meeting can be replicated due to their defined steps and adherence to established methodologies. The EU provides general principles for organising these groups in the Methodological material for CSP preparation, aiming to ensure results' replicability. However, the actual replicability varies, since there are diverse national methodologies that extend beyond EU guidelines, regarding desk research, document analysis, and adherence to country-specific internal policy documents. These national methods differ significantly across the EU, affecting the organisation and outcomes of focus group discussions. Most respondents also reported that Focus groups often present some limitations such as transparency issues due to insufficient documentation, language barriers of protocol produced at national level not allowing external review, lack of anonymity, which could discourage participation, especially from hard-to-reach and distant populations, and the limited number of participants they can accommodate.

***Applicability**: high*, these tools *are easy to adapt to different context and policy steps, contribute to SO1, 4 and 7, and some MS already implemented them in the previous programming period.*

Focus groups and meeting results are easily adaptable to different contexts, like marketing and political science, addressing different research questions relevant to policymaking. Thanks to this multifunctionality and versatility, MS widely applied these across different governance levels in identifying local needs, informing regional and national policy, and shaping EU priorities. In terms of policymaking steps, the respondents of the survey is reported that focus groups have been used mostly in the design of interventions, needs assessment and SWOT analysis, but also for socio-economic analysis and budget allocation, and a higher number of responses suggest that they have especially contributed to SO1, SO4 and SO7 of the CAP. In the previous CAP programming period, focus groups and meetings were organised in many EU MS[35], with differences in organisation, practices adopted, and responsible government bodies, with distinction between external and internal focus groups.

***Accessibility**: high, these tools are easy to use with limited knowledge and no training requirements, although guidelines are generally available, survey respondents are not always aware of their existence.*

Focus groups and meetings are easy to arrange, and no special skills or training are formally required. This view is acknowledged by 57% of survey respondents, while 43% noted the necessity of trained specialists for effective policymaking. Enhanced trainings could improve their use in CSPs design and monitoring. Only a few MS, including Belgium-Flanders, Ireland, and the Netherlands are currently applying focus groups and meetings tools with expected proficiency. Despite the existence of various technical and methodological guidelines related to the implementation of focus groups and based on scientific research or policy consultations, many respondents of the

35 Including Belgium-Flanders, Cyprus, Ireland, France, Lithuania, Bulgaria, Slovakia, Spain, and the Netherlands.

survey reported that are often unaware of them, indicating a need for better awareness and utilisation of these resources.

***Efficiency**:* high, implementation *time and effort vary, but the cost is often low compared to other involvement methods due to the limited number of participants required.*

The time required for implementing Focus Groups and meetings varies significantly. Organising them demands a considerable commitment, especially when it involves analysing and summarising discussions. According to 60% of survey respondents, one month is necessary to set up a focus group, while about 30% report that more than a month may be needed. The overall cost can be very low compared to other stakeholder involvement methods making it a cost-effective way to gather information from specific target audiences due to the limited number of people needed. The cost of a focus group depends on various factors, but it typically ranges from 3,500 to 4,500 Euro for a single, professionally conducted session (Gell, 2023). However, 43% of survey respondents indicated that often less than 1,000 Euro are needed for the organisation. The number of participants in focus groups and meetings varies based on several factors such as the number and length of sessions and the incentives provided (Gell, 2023). Standard group sizes range from 4 in smaller groups to 12 in larger groups. According to 43% of respondents, less than 5 people are necessary for implementation, but 29% reported that 5-10 people are required.

Analysis by criteria

ACCURACY

Accuracy analysis considers how successful is the tool 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: Are measures in place to avoid mistakes and redundancy and to ensure consistency in the stakeholder's responses or information provided (e.g. to verify the quality and reliability of the information)? Are results triangulated, corroborated, or substantiated by other data sources? Are different types of stakeholders involved?

Data input and its sources are well traceable and validated, since the content of discussions, performed in focus groups and meetings, are normally recorded and transcribed in protocols in all MSs. There is a difference in the content and accuracy and availability of protocols. Some MSs collect and publicly store in websites very detailed protocols from the focus groups discussions. For example, in Cyprus, the outputs from focus groups include detailed discussion notes that provide a basis for supporting interventions and setting targets in agri-environmental measures. In Ireland, the outcomes of focus groups and meetings are compiled into detailed written reports, available on the Ministry's website, and the reports summarise the discussions, ideas, and consensus reached, serving as reference materials for policy formulation and decision-making[36]. The same practice is also applied in Lithuania. In Romania, individually crafted questionnaires for each topic's working groups are stored in a database containing the responses, enabling a detailed understanding of stakeholder preferences and views. In Spain there are also developed detailed reports from focus groups and discussions[37,38,39], that summarise stakeholders' reactions to the topics presented, reflecting the diverse viewpoints and contributing to a more comprehensive understanding of the subject matter.

[36] www.gov.ie/en/publication/76026-common-agricultural-policy-cap-post-2020/#cap-reform-public-consultations-and-public-engagements

[37] https://www.mapa.gob.es/es/pac/pac-2023-2027/analisisjornadapac_tcm30-525185.pdf

[38] https://www.mapa.gob.es/es/pac/pac-2023-2027/informederesultadosdelprocesodeparticipaciondelatercerajornadadepartenariado_tcm30-556528.pdf

[39] https://www.redruralnacional.es/documents/10182/805445/Informe_final_4Reunion_Partenariado.pdf

According to our survey, it is not clear to most respondents whether measures were in place to avoid mistakes and redundancy in the stakeholder's responses. Also, it is difficult to generalise regarding the triangulation, corroboration, or substantiation of focus group discussion results by other data sources. For instance, in the Lithuanian case, this is always a question of time available for a particular topic: at the beginning, all possible other data sources are used for initial discussion phase (CSP Design); however, in further stages time limits force to jump through the use of other available data sources.

A common practice of all MSs is to involve a range of different types of stakeholders in CSPs preparation. Due to the huge number of stakeholders to be involved (considering the method), they are normally allocated into thematic focus groups (in many cases these are also called Thematic Working Groups) according to different topics. For example, Belgium-Flanders form 5 working groups: 1) Environment & Climate; 2) Economic Resilience; 3) Rural; 4) Innovation & Knowledge, and 5) Farmer definition. The same practice is applied in other countries, only the topics and number of them differs. Partnership meetings involve the participation of representatives of all relevant actors (public and private) who are directly and/or indirectly affected by CSPs and aim, for example in Spain, to collect feedback from participants regarding Socio-economic analysis, SWOT analysis, Needs identification and prioritisation.

According to a focus group from Poland, the highest score of accuracy of tools was given to ''focus groups and meetings'' with a score of 4.4 (over 5). However, a Lithuanian focus group said that the accuracy of focus groups tools can vary since it can be used as a political tool and a lot depends on the composition of the tool. According to them, the focus groups tool is good to collect different opinions but also it can be very political as every member of the focus group represents certain interests that he seeks to represent. According to Italian focus groups, for online questionnaires, there is a validation step through peer review and comparison with results of other tools. While for the consultation with specific communities, there is a validation step through validation by external experts.

Tool Representativeness

Representativeness refers to 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: How many stakeholders can be involved in the implementation of this sub-category of tools (e.g. number of participants of the workshop, focus group, consultation etc.)? Does this sub-category of tools facilitate the involvement of all relevant stakeholder groups while ensuring representativeness? Is this sub-category of tools affected by biases?

According to our survey, focus groups and meeting tools typically involve a modest number of stakeholders, ranging from 10 to 25, as observed in the focus groups implemented in the Slovak Republic, which engaged stakeholders at the national level, and in Romania, which involved regional and local stakeholders for the design of interventions. However, in Lithuania, where several focus groups were held, the number of participants ranged from 50 to 250 stakeholders at the national, regional, and sub-regional levels. Most of survey respondents indicate that the tool engaged stakeholders from different governance levels.

Focus groups hold high capacity as a tool to accurately and comprehensively capture or depict a system, dataset, or phenomenon. For example, in Ireland, Town Hall Meetings have been implemented to foster an inclusive and participatory environment, allowing for a broad spectrum of views from various regions and stakeholders. This approach seeks to engage the public, industry representatives, and policymakers in open dialogue and collaboration on agricultural policies. Other MSs (e.g., Belgium-Flanders, Cyprus) highlight that they utilise internal particular Ministry's/Department's expert groups (titled CAP Working Groups or similar) to discuss the proposals for interventions within the different themes and identify needs accurately, leveraging their expertise in the SWOT analysis and structured decision-making. They are beneficial in driving consensus decisions on key policy issues related to CSP design. Slovakia convened 21 sub-working groups (the Ministry of Agriculture and Rural Development of the Slovakian Republic) to finalise intervention wording, drawing from a larger Working Group.

There also exist other practices, when issue-specific experts, who represent the united/summary position of particular interest groups (not their own professional opinions) are invited to take part in focus group discussions and meetings, to foster an innovative view on the development pathway of particular fields. For instance, in Lithuania, the Ministries of Agriculture and Environment normally apply participatory approach by involving diverse stakeholder

representatives from different sectors, i.e., science, policy, business, and topic-specific NGOs. The Spanish National Ministry of Agriculture utilised Focus Working Groups to involve diverse specialists from different areas, who may not be directly related to agricultural policy, in the analysis of the agricultural sector to foster the innovative trends and perspectives.

It was mentioned by interviewees that sometimes there is risk to prompt biases-affected opinions in focus groups, when particular lobbying interests are involved. However, it is up to professional running of discussion to avoid such practices. The level of representativeness in the partnership meeting is high because this tool, as it was used in Spain, involved the representatives of the most relevant stakeholders (Ministries, Governmental bodies, Regional and Local Authorities, Research and Scientific institutes, Consulting firms and companies, Farmers and Agricultural organisations, International Organisations, Environmental and consumer organisations), potentially interested in the impact of the NSP.

Summing up, focus groups as a tool holds capacity to serve accuracy and coverage parameters from several perspectives: first, to ensure high coverage, when calling for public discussions regarding particular issue and then splitting into thematically focused sub-groups; and second, by giving in-depth accuracy via cumulative representative expertise (e.g., NGOs, such as National Family Farmer's Associatiotion, etc.), when working in expert-based focus groups.

According to the Polish focus group results, the number of participating stakeholders is directly related to the type of tool applied. Focus groups and meetings involved on average the lowest number of participants – up to 20 people. However some higher numbers were observed in other countries as Germany, where the number of stakeholder groups involved in the national accompanying committee was 45. According to a Romanian focus group, working groups are not very rigorous because participants have different backgrounds and hence different levels of understanding the context and ability to get an overall view. Some of them had limited views, while others had extensive ones. On the other hand, all participants in the German focus group on the ''accompanying committee'' tool saw the procedure of expression of interest for the stakeholder consultation that was implemented to select the stakeholder groups for the monitoring committee as a success in involving all relevant stakeholders. Subsequently, interest groups were founded that represent several organisations and are open to new members.

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 consistency and stability of the results obtained from the application of focus groups and meeting tools are diverse. Focus groups have historically been employed across a multitude of disciplines to research and solve a variety of challenges including generating hypotheses, exploring opinions and attributes, and developing new product ideas (Fern, 1982; Hair et al., 2008; Kitzinger, 1995; Murgado-Armenteros et al., 2012; D. W. Stewart & Shamdasani, 2014). Focus groups are not without their limitations though, as they lack anonymity, discourage participation from hard to reach and distant populations, are limited in the number of participants they can accommodate, and can be relatively expensive to conduct (Gammie et al., 2017; Prescott et al., 2016; Richard et al., 2021).

Concerning the CAP SP preparation, from one point of view, the replicability of Focus group and other meetings' discussions depends on the comprehensive methodology applied via its conduction process. However, the desk research and document analysis reveal that the Focus groups methodologies in the CSP preparation processes highly differ across MSs. There exist different ministry's methodologies (normally - internal policy documents)

concerning the exact organisation of focus groups in a particular country. From the other side, the general guiding principles on focus groups and meetings organisation are clearly stated by the EU in methodological material for the CSP preparation, and it is up to the national-level focus group discussion organisers to follow this advice to make the results replicable. According to our survey's respondents, in their countries (Slovakia, Portugal and Romania) focus groups were implemented following established methodologies.

According to the Polish focus group, for consistency of results, it is necessary to be able to identify the respondent, and this is best fulfilled by focus groups and face-to-face meetings. Similarly, a Romanian working group said that the stakeholder category to which the respondent belongs can affect the coherence of answers. In addition, they also indicated that the participants expertise and farm size can also affect it.

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?

The availability and accessibility of information concerning focus group methodology, data and results are diverse in MSs. According to our survey, all seven respondents stated that, as far as they know, there is no available documentation on the underlying methodology. For example, in Cyprus and Bulgaria, there is no publicly available documentation concerning the focus groups discussions. The protocols from focus groups' discussions are normally written in national languages, and not always open to external review. It is mostly a subject to the professional proper documentation of focus group discussions, which depends on the organisation authorities. For instance, in Ireland the outcomes of focus groups meetings are compiled into detailed written reports, available on the Ministry website[40], and these reports summarise the discussions, ideas, and consensus reached, serving as reference materials for policy formulation and decision-making. In the Lithuanian case, there are publicly available data from focus groups discussions[41], that consist of detailed SWOT analysis reports, summarising the insights and prioritised needs identified through stakeholder engagement. Detailed reports that summarise stakeholders' reactions to the topics presented, reflecting the diverse viewpoints and contributing to a more comprehensive understanding of the subject matter were elaborated in Spain, after each of four partnership meetings[42].

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?

[40] www.gov.ie/en/publication/76026-common-agricultural-policy-cap-post-2020/#cap-reform-public-consultations-and-public-engagements, https://www.gov.ie/en/collection/79383-cap-post-2020-public-submissions/

[41] https://zum.lrv.lt/lt/lietuvos-zemes-ukio-ir-kaimo-pletros-2023-2027-m-strateginis-planas-1

[42] https://www.mapa.gob.es/es/pac/pac-2023-2027/analisisjornadapac_tcm30-525185.pdf, https://www.mapa.gob.es/es/pac/pac-2023-2027/informederesultadosdelprocesodeparticipaciondelatercerajornadadepartenariado_tcm30-556528.pdf, https://www.redruralnacional.es/documents/10182/805445/Informe_final_4Reunion_Partenariado.pdf

The use of focus groups has spread widely in marketing studies since the 1970s, however, they are less frequently used in political science (Berkhout et al., 2023). This is surprising because a wide variety of research questions in policymaking addresses the behaviour and interactions between such actors, which mirror positions, ideas, and power relationships. Focus groups offer a way to observe such interactions in a purposefully designed setting, varying, for instance, the topics, cases, or scenarios to be discussed, the composition of the focus groups, or the role of the moderator (Cyr, 2016).

In contrast to the above evidence from the literature, MSs widely used focus groups and meetings in CAP policymaking at different steps and levels for a long period of time, especially due to the multifunctionality of this tool. Focus groups discussions can be used at all policy-making steps to pursue different objectives at the same time. All MSs define relevance and multi-functionality of the focus groups and meeting tools throughout all CAP objectives and steps of policymaking process, including design and monitoring. According to the survey respondents, this sub-category of tools contributed to the discussion regarding all CAP specific objectives, and to inform all policy tasks (and especially for interventions setting, needs assessment, socio-economic context and SWOT analysis, and Budget allocation). For example, Cyprus refer to another relevant practice of particular external focus groups, that are used in the process of specific objectives 4, 5 and 6 for formulating CAP interventions in the field of Agro-Environmental and climate measures (Art. 70). In Spain the partnership meeting tool was used for collecting of all stakeholders' opinions on socio-economic analysis, SWOT analysis, needs identification, and prioritisation. Implemented practices in MSs elucidates that focus groups can be easily tailored or re-adjusted to assess or contribute to the assessment of the other Specific Objectives and policymaking steps, depending on the existing needs.

Our focus group results demonstrated that focus groups and meetings tools are highly multi-functional. The Lithuanian focus group indicated that the tools can be adapted to different contexts as Lithuania, when preparing the SP, ten focus groups were held according to nine specific objects and one horizontal objective related to AKIS. In addition, in Poland, focus groups and meetings were considered the most flexible and adaptable tools, with the highest score of 4.2.

Experience

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

Evaluation Question: How experienced are MS 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?

Focus groups and meetings have been widely applied in MSs for Common Agricultural Policy (CAP) making practice, especially when the policy course of the EU shifted towards the transformative policies for the solidarity, green and social economy (Chaves-Avila & Gallego-Bono, 2020; Lovrić et al., 2023). According to our survey, in the previous CAP programming period, focus groups and meetings have been organised in many MSs, including Belgium-Flanders, Cyprus, Ireland, France, Lithuania, Bulgaria, Slovakia, Spain, and Netherlands, among others. For the current CSPs, as many MSs have used focus groups and meetings, including Slovak Republic, Portugal, Romania, Slovenia, Spain, and Lithuania, at least.

There are different focus groups and meetings organisation practices across MSs. First, some MSs highlight the distinction made between external and internal focus groups (e.g., Cyprus, France, Ireland, Lithuania, Spain), which act in several rounds. External focus groups as a form of public consultation involve diverse specialists and the general public actors from different areas (business, science, society), who may not be directly related to agricultural policy. They are important in the general analysis of the agricultural sector and its possible development trends. This initiative arise normally from the responsible government bodies, except France, where this is done by Scientific and Research Institutes and Representatives of civil society. Internal focus groups are formed in most cases by Ministries of Agriculture or equivalent to finalise intervention wording, drawing from a larger focus group (e.g., Ireland, Lithuania, Slovakia). Most of focus groups and meetings in MSs are performed in several rounds (except Belgium-Flanders).

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? What is the level of detail that these tools can reach?

Focus groups and meetings might well support the analysis and policymaking at different levels, including local, regional, national and EU level. For example, In Italy's focus group about consultations with scientific communities, it was suggested that these tools can be used at EU level. MSs apply this tool at all levels, depending on the objectives and tasks. For example, in Spain the partnership meeting tool involve all spectrum of actors at different level of representation (regional, national and international) and, through that, the tool could provide a good balance and effective dialogue among all governance layers. In Spain, focus groups are organised for each SO during the CSP design phase. First, regional Ministry's departments use this tool at regional level, then – national Ministry's departments use the generated results further in CSP Design process. Thus stakeholders from Ministries' Governmental bodies, Regional and Local Authorities, Research and Scientific institutes, Consulting firms and companies, Farmers and Agricultural organisations, International Organisations, Environmental and consumer organisations take part in CSP design process at different level focus groups.

At local level, thematic focus groups and meetings are utilised to envisage the territory-relevant or issue-specific needs (see for example SHERPA Rural Interfaces[43]). In turn, these needs are reflected in regional and/or national level focus groups (several of them), which generate the proposals for interventions considering the CAP specific objectives. Further the EU-level focus group might compose the overall strategic objectives for the EU CAP policy.

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?

Focus groups are very useful, because the discussion among participants will help to clarify opinions, provoke more in-depth reasoning, and to disclose whether opinions are shared by many. Whilst a focus group discussion is a qualitative research approach, it also enables a semi-quantitative analysis of statements made (Bergen, N., & Labonté; Hennink & Kaiser, 2022).

Among the policymakers and elsewhere exist a commonly accepted view, that focus groups and meetings are easy to arrange, and no special skills or training is needed. Accordingly, when asked about their satisfaction with the tool in terms of ease of use, survey respondents scored on average 3.7 out of 5. Four respondents out of seven think that no prior training is required, but three respondents suggest that a training up to 20 hours may be beneficial. However, it should be noted, that in policymaking process, focus groups and meetings should be arranged and organised professionally, by trained specialists, and be well documented (Busetto et al., 2020), in order to fulfil their targets in policymaking.

According to the results of our focus groups, the ease of use was assessed differently, sometimes even by different stakeholders within the same focus group, as was the case in Romania. While in the case of the Slovakian focus group, from the perspective of the managing authority, these tools were the easiest to use and provided what was

[43] https://rural-interfaces.eu/rural-interfaces/

needed as an input into interventions formulation and setting. On the other hand, the stakeholders of this Slovakian focus group that instead of instead of simplification there is a tendency to complicate things even more.

Tool Availability

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

Several guidelines are available for the implementation of focus groups, and it is a well-know and long-standing tool in science and public consultations.

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?

A focus group is a research method that brings together a small group of people to answer questions in a moderated setting. The group is chosen due to predefined demographic traits, and the questions are designed to shed light on a topic of interest. For example, Liamputtong P. (2011) describes focus groups methodology on many aspects, such as theory, ethics, methods, methodology and many other.

Therefore, there are several technical and methodological guidelines available based on examples from scientific research or policy consultations and there many platforms and researchers analysing the method of focus groups, as well as there are different sampling methods. As mentioned above, according to our interviews and survey, several MSs have used focus groups for current and previous CAP programming period, thus there is string experience with the tool.

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 sub-category of tools (i.e time necessary to organise and prepare the implementation, including the identification and invitation of participants, setting up the online page of the consultation or - in case of methods implemented in presence - the arrangement of the venue, etc.)? And how much time does it take to implement the process of the tool (i.e. time necessary to implement the tool until the achievement of the intended outputs)?

Focus group and meetings tools can be difficult to organise as they require commitment from stakeholders and important organisational resources. To hold a focus group, it is necessary to organise and prepare the implementation, including the identification and invitation of participants, setting up the online page of the consultation or - in case of methods implemented in presence - the arrangement of the discussion itself, input materials, prompts, venue and others. Therefore, often the focus group implementation is difficult, and it can take a big amount of time to organise, to hold and to analyse or summarise the discussions.

Also, 60% respondents to the survey declared that up to 1 month time is necessary for the setting-up a focus group and around 30% reported that more than 1 month can be needed. This becomes even more difficult when the same organisation, e.g. Ministry of Agriculture (or other public organisations) has to organise an extensive amount of focus groups, for example organising focus groups for all CAP objectives. When it comes to the implementation of these tools, 43% of respondents reported that between a day and a week are sufficient, while the remaining 57% consider that at least one month of implementation is necessary, with some respondents (29%) declaring that more than 6 months are required. Accordingly, in our focus group results, there was tendency to say that the focus groups and

meetings required a significant amount of time to organise. For example Romania's focus group indicated that more than 6 months were needed for the organisation of interviews with farmers and focus groups.

Cost-effectiveness

Cost-effectiveness refers to the extent to which the tool is able to achieve its outcomes within a reasonable amount of 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 budget can be necessary for the maintenance of the tool or subscription for patent-based tools every year?

The cost of a focus group depends on several factors, but as a general rule, it is expected to pay on average about 3,500 - 4,500 Euro for a professionally ran single focus group at the very end (Gell, 2023). When asked about the implementation costs, 43% of our survey respondents declared that these tools require less than 1,000 Euro for their implementation. Nevertheless, a share of 14% reported that more than 5,000 Euro are necessary. According to a Polish focus group, focus groups and meetings are less resource demanding than events involving more stakeholders (workshops and conferences), but certainly more time and financial resources are needed to organise and conduct them.

It is worth noting that this is the cost per one focus group or meeting. Normally, it is expected to run multiple focus groups (given a focus group size of 8–12 people) aiming to get a good range of data (Hennink & Kaiser, 2022). Because focus groups are qualitative instead of quantitative research, it is very easy for a single focus group to give a skewed perception of whatever it is under research/consensus — that is why professionals recommend to run at least three or four focus groups (Busetto et al, 2020). Moderator can cost between 650-1,300 Euro per one focus group. The aspects to consider before planning the budget to run focus groups: number of participants; number of focus groups; length of the focus groups; incentives; location of the focus groups (live or virtual).

Human Resources Efficiency

Human resources efficiency refers to the extent to which the tool is able to achieve its outcomes involving a reasonable amount 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?

Methodological literature suggests that the amount of people, necessary for the implementation of the tool and the analysis of results, vary according to the number of participants, number of focus groups and meetings, the length of focus groups and the incentives of them (Gell, 2023). Standard sizes of groups can range from as little as four mini-groups to 12 large groups. Smaller groups allow to have more one-to-one conversations to each participant. Large groups require a professional moderator to feed discussions (Hennink & Kaiser, 2022). According to 43% of our survey respondents, less than 5 people are necessary to implement the tool. However, 29% reported that 5-10 people are required while 14% consider between 25 and 50

There are two main professionality – moderator and data analysis specialist, that are necessary for the good implementation of the standard size (up to 12 people) tool (Busseto, 2020; Lobe, 2020; Lochmiller, 2021). The larger the group is, the duration of results' analysis should be taken into account. The same analyst might work longer for the larger focus group and meetings' results, however timescale for policymaking steps matters. Professional analysis of results from a standard focus group or meeting group might be implemented by 1-2 people – analysts, who took part in the focus group process (Hennink & Kaiser, 2022). The professional is needed in order to follow the requirements for data analysis (Busseto, 2020; Lochmiller, 2021).

3.1.5. Key findings

In this section, we summarise and cast light on the key aspects emerging from the evaluation of stakeholder needs assessment tools, and we conclude by identifying the strengths and weaknesses of this category of tools.

Stakeholder needs assessment tools include all those tools employed to gather opinions, information and feedback from experts, stakeholders, and the wider society to inform the design of the CSPs. They have been mainly used to inform the first design steps such as the socio-economic context, SWOT analysis, and the needs assessment, but also to provide feedback on the draft CSPs.

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

Accuracy is assessed as limited to moderate (online consultations, workshops and conferences) or moderate (focus groups and meetings). In particular, online consultation, survey tools, and workshops and conferences have more limited accuracy due to issues related to stakeholder selection, reliance on stakeholders' perceptions, lack of traceability of input data, and self-selection biases. Focus groups and meetings can allow a slightly higher accuracy because they often involve expert knowledge, methodological protocols, and input data can be more easily validated.

Reliability is judged as limited (workshops and conferences) or moderate (online consultations, focus groups and meetings). Online consultation tools and surveys as well as focus groups and meetings were easy to replicate, but sometimes resulted in varying outcomes due to different factors, such as sampling bias and self-selection for online consultations and surveys. On the other hand, workshops and conferences were less replicable since the methodologies are not implemented systematically and not documented.

Applicability is assessed as high for all the sub-categories of tools. The three sub-categories of tools were considered flexible as they were easy to adapt to different contexts, policy steps, and CAP SOs. In addition, Managing Authorities have long-standing experience with them since they have been widely used in previous programming periods, especially workshops and conferences,, focus groups and meetings.

Accessibility is also judged as high for the three sub-categories of tools since they are widely and openly available. All tools were reportedly easy to use with limited previous knowledge, such as survey design, technical skills for online consultations and surveys, and specific facilitator skillsets for workshops, focus groups, and meetings.

Efficiency is assessed as moderate (workshops and conferences) or high (online consultations, focus groups and meetings). Online consultations and surveys were highly efficient due to the diverse audience they could potentially reach, the quick implementation time, and their reduced cost. However, the implementation time could be long. Similarly, focus groups and meetings were highly efficient due to the limited number of participants. On the other hand, workshops and conferences were moderately efficient due to their relatively higher cost and the more extended preparation periods needed.

Stakeholder needs assessment tools play a crucial role in shaping policy design. While these tools exhibit high applicability and accessibility, offering flexibility across various policy steps and objectives, they also present notable challenges. The limited to moderate accuracy and replicability of online consultations and surveys and the less systematic approach in workshops and conferences pose some concerns. These concerns particularly relate to the transparency and potential integration of the results into the CSP design process, as well as to the potential stakeholder fatigue perceived during consultation activities, which could influence the variability of the obtained results. Despite their efficiency, the extended preparation periods and higher costs associated with workshops and conferences warrant consideration.

In conclusion, stakeholder needs assessment tools prove valuable, especially in the early design steps and feedback processes, yet careful attention to their limitations is essential. The limited reliability and accuracy of these tools might suggest that they are less suitable for informing highly technical tasks like intervention or target setting, for which other tools like policy analysis might fit better. Complementing stakeholder needs assessment tools with more evidence-based tools (like policy analysis tools) or structured tools (like policy choice supporting tools) might help fill the accuracy and reliability gap.

3.2. Policy choices supporting tools

3.2.1. Inventory of policy choices supporting tools

Table 5 below shows examples of tools belonging to the different sub-categories, according to our inventory of tools. The full list of tools is available in the Tools4CAP online inventory.

Table 5. Sub-categories of tools and examples of tools from the Tools4CAP inventory.

Sub-category of tools	Example of tools identified
Voting and prioritisation tools	• Constrained Cumulative Voting (Italy) • Cumulative voting approach (Lithuania) • Voting for Needs Prioritisation (Romania) • Needs Scoring Tool (Slovenia) • Number-Voting (Germany)
Multicriteria analysis	• Multicriteria Analysis (Spain) • Logic Model (Romania)

Policy choices supporting tools have been employed to support different tasks of the CSP design and monitoring. Table 6 below provides an overview of the number of MSs where a specific sub-category of tools was employed for each task of the CSP design and monitoring. More details are available in the online inventory of tools[44].

Table 6. Use of policy choices supporting tools across the CSP design and monitoring (no. of MSs in which the tools were employed for the specific task).

	Voting and prioritisation tools	Multicriteria analysis
CSP design tasks		
Socio-economic context analysis		
SWOT analysis		
Needs assessment	5	1
Interventions setting		
Target setting		
Financial allocations		1
Ex-ante analysis and SEA		
Stakeholders' consultations		
CSP monitoring tasks		
Performance review		
Beneficiaries' compliance		
Evaluation		

The following sections presents the results of the evaluation specific for each sub-category of tools.

[44] https://www.tools4cap.eu/tools/

3.2.2. Voting and prioritisation tools

Summary

***Accuracy**: limited to moderate. The tendency of these tools to simplify complex issues negatively affects result accuracy due to cognitive effort, data quality, and subjective interpretations. Despite limitations, its structured approach fosters shared consensus and objectivity.*

Voting and prioritisation (VP) tools offer structured approaches to prioritise needs but face challenges such as oversimplification and susceptibility to strategic behaviours. Combining VP with qualitative assessments enhances holistic representation. VP yields clear documentation for policy guidance, however there are still drawbacks such as prioritisation techniques and vote splitting. Voting for needs prioritisation can benefit from data-driven approaches, leading to more accurate and valid decisions. However, limitations in data quality, subjective interpretations, and incomplete information can potentially undermine the accuracy and validity of the prioritisation process. VP provides a quantifiable representation of needs, enhancing decision-making and resource allocation with relative objectivity. VP encourages high stakeholder participation and cluster results, facilitating collective consensus in large groups. While useful for reflecting diverse preferences, careful wording and consideration of options are vital to maintain representativeness, emphasising the need for supplementary stakeholder consultation.

***Reliability**: limited to moderate, because complexity, subjectivity in allocation, and reliance on specific contexts limit the tool's replicability. While other approaches might not replicate VP directly, they could aim to achieve similar outcomes by allowing participants to express weighted preferences. However, replicating the exact distribution of votes in cumulative voting might be challenging due to its intricacies.*

VP is a scalable and accessible method for various needs but may face challenges in replicability due to its rigid structure and potential subjective variations. VP, considered fast and accurate, requires clear understanding for replicability, with challenges arising from its complexity and context dependence. While achieving precise replication might be difficult, alternative methods could strive for similar outcomes by accommodating weighted preferences or allocations, acknowledging the intricacies of VP. Assigning numerical values in number voting enhances transparency in prioritisation, reducing decision-making ambiguity. However, unclear criteria for scores can compromise transparency, necessitating clear communication and detailed criteria explanations. Integrating qualitative insights and flexibility enhances transparency.

***Applicability**: moderate to high, but the use of these tools is currently limited to few MSs. They can support at different levels depending on their design. No specific evidence of limitations to their applicability in other context - More research is needed to assess their potential.*

Voting and prioritisation tools, when used for needs assessment, have limited multi-functionality in addressing various policy objectives. They are mainly used to facilitate collective discussions and prioritise identified needs. Although adaptable to different CAP objectives, modifying these tools for diverse policymaking steps is restricted. VP, while applicable at all administrative levels, faces constraints in introducing new options once initiated. Only five MSs stated the use of this tool in the preparation of their CSPs. When it comes to granularity, the analysed sub-category of tools can facilitate analysis and policymaking at various administrative levels. The tools' ability to reach specific levels of detail is contingent on their design, including the scoring range and the selection of experts involved in the prioritisation process.

***Accessibility**: moderate to high, as they are generally easy to use if a good methodology is in place (and therefore designed) voters also needs adequate education and understanding of the system to be able to use it effectively. These tools are available and accessible to everyone.*

The tools are generally user-friendly, requiring minimal training time. Although often implemented by managing authorities, external expertise may be needed for methodology design. Most VP methods are easily accessible, accommodating users of varying expertise. Other VP approaches, regarded as fast and accurate, require adequate education for effective utilisation, emphasising the importance of understanding the system for optimal use. These tools are usually freely available, but tailoring some methodologies entails costs. The main challenge lies in designing context-specific methodologies, with complexity in determining resource needs. Stakeholders need

adequate education and understanding of the system to utilise it effectively, which might require additional resources for outreach and education.

***Efficiency**: moderate to high, as time and costs are moderate although some expertise is needed when it comes to staff involved.*

Time efficiency of these tools depends on their design. Interviewees mentioned that the tools were tailored to the resources available, including time. When it comes to costs, it could vary depending on the levels of complexity, if it is internally designed or not, the number of steps required to reach the hierarchy of needs, number of experts and their associated costs, and scope of the task to be fulfilled (how many CAP priorities it includes). Similarly, the number of people necessary to implement this tool depends on the scope of use of the tool and the complexity of it as well. Additionally, this tool requires a certain level of skills from the people involved in the tool design, implementation, and analysis of the results.

Analysis by criteria

ACCURACY

Accuracy analysis considers how successful is the tool 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: 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? Is it possible to easily trace back the information used for supporting the policy choices to its original source? 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)?

A drawback that is common to most prioritisation techniques is that the technique becomes harder to use with many items and stakeholders may lose perspective (cognitive effort). In addition, the basic form of voting tools is susceptible to 'shrewd tactics' which may arise when stakeholders distribute their points based on what they think others will do to raise the priority of their favourite items. Finally vote splitting might arise for similar or related options. In Lithuania, this tool resulted in clear documentation of prioritised needs and intervention settings, providing structured guidance for policy formulation. Sahni (2007) also conducted a controlled experiment on cumulative voting and found that it was a scalable and accurate method but noted that ease of use and learning could be potential challenges.

Voting for needs prioritisation can benefit from data-driven approaches, leading to more accurate and valid decisions when supported by reliable and comprehensive data (Cagliero et al., 2022). However, limitations in data quality, subjective interpretations, and incomplete information can potentially undermine the accuracy and validity of the prioritisation process. Besides, their tendency to oversimplify complex issues may overlook qualitative aspects and context. Thus, pairing voting tools with qualitative assessments ensures a more holistic representation of needs in prioritisation. This approach was used by the Romanian Ministry of Agriculture to leverage the internal expertise who had prior experience and good knowledge of EU rules, objectives, and priorities. It also allowed for efficient needs prioritisation within time constraints.

As for number voting, numeric values provide a quantifiable representation of needs, making it easier to compare and analyse different needs against each other. This quantification aids in decision-making and resource allocation. Though no method is bias-free, number voting tends to be more objective due to reliance on numerical values, fostering a balanced representation of needs.

From the Italian focus group results it emerges that, overall, the proposed tools are quite accurate and guarantee a homogeneous representation of participants. Balancing traceability and privacy are pivotal in any voting process,

ensuring robustness against fraud while safeguarding individual privacy. Achieving this delicate balance is a challenge. Despite an abundance of guidelines addressing these concerns, there's a notable lack of information specifically pertaining to the mentioned voting tools. This gap emphasises the need for tailored guidance in implementing traceable yet privacy-preserving systems within these contexts.

Special attention should be given to the wording of the different options put to the vote and/or the definition of the criterions used for selecting the different options, since this can clearly influence the voter and therefore ultimately lose the representativeness of the opinions that would have been expressed in an open question context. In addition, it is important to underline that it is not the only way of expressing levels of priority and that it must be accompanied by adequate stakeholder consultation (Cagliero et al., 2021).

Furthermore, respondents to our online survey indicated that there are no specific measures in place to avoid mistakes and redundancy in the input data. On average, they rated the tool as based on well-founded evidence and the information source (e.g., documents collected) clearly identified and documented with a 3 out of 5. Nevertheless, in the focus group conducted in Italy it was highlighted that the results obtained are validated through peer review and comparison with results of other tools. Similar outcomes were observed also in the Slovenian case, where the validation of the evidences obtained from the needs prioritisation approach takes place at several levels, involving incorporating feedback from drafters and the public consultation into tool development, with adjustments made as needed based on this feedback, until the CSP is finalised.

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: Are the tools able to capture the diverse facets, variations, or elements within the system or dataset it seeks to represent? Does this sub-category of tools facilitate the involvement of all relevant stakeholder groups while ensuring geographical and sectoral representativeness?

Voting tools can be applied to a wide range of needs and situations, making it a scalable method for representation, allowing the collection and integration of stakeholders' perspectives. This has also been confirmed during the Italian focus group, where participants highlighted how overall the proposed tools guaranteed a homogeneous representation of participants. However, the two survey respondents answering for this sub-category of tools indicated that these tools have not been used to capture different governance level, although potentially they could be adapted to this purpose.

Voting tools allow for a high rate of participation among stakeholders and the possibility of clustering results, rather than merely providing a list of priority values (Tufail et al., 2019). As such, it is particularly useful for reaching a collective consensus with a large group of participants and a high risk of disagreement. It can lead to outcomes that better reflect the spectrum of preferences within a group or electorate, rather than simply favouring the most popular choice (Cagliero et al., 2022). In Italy, constrained cumulative voting was introduced to facilitate collective discussion and prioritisation of the territorial needs. The process generates a shared consensus on the importance of each need and categorises them into homogeneous groups according to their importance for intervention.

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?

While replicating the results of voting tools is feasible; it might face challenges due to the subjective nature of assigning scores and potential variations in interpretation among different individuals or groups. Adjustments may be necessary to align with specific contexts or to accommodate differing interpretations while aiming for similar outcomes. Replicating the results of voting for needs prioritisation faces similar challenges. Cumulative voting was identified as being one of the fastest and most accurate methods that our survey respondents also rated as easy to use. It also allows voting with fine-grained information on voter preference intensity (Cagliero, 2021).

However, ensuring that voters have a clear understanding of the system is necessary to ensure the replicability of the initiative as well as the quality of the data, as additional resources for outreach and education might be required. In essence, replicating cumulative voting results precisely might be difficult due to its complexity, subjectivity in allocation, and reliance on specific contexts. While other approaches might not replicate cumulative voting directly, they could aim to achieve similar outcomes by allowing participants to express weighted preferences or allocations through alternative voting methods. However, replicating the exact distribution of votes in cumulative voting might be challenging due to its intricacies.

The Italian focus group indicates that the cumulative voting approach has been replicated on a regional and local scale. Conversely, during the Slovenian focus group, the replicability of the Deloitte/AIS report was assessed as quite low, primarily due to the timing of the report's preparation, which preceded the adoption of the final CAP legislative acts. This necessitates additional work after publication, resulting in a significantly different categorisation and, consequently, prioritisation of needs.

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?

Assigning numerical values to needs creates a clear and quantifiable prioritisation method, fostering transparency in comparing and ranking them. This structured approach reduces ambiguity in decision-making, minimizing misunderstandings. However, lack of clarity in the criteria for assigning scores can lead to inconsistencies and compromise the validity of results, impacting transparency. To mitigate these issues, supplementing voting tools with clear communication, detailed criteria explanations, and open discussions is crucial. Integrating qualitative insights and flexibility in prioritisation enhances transparency and ensures a comprehensive representation of needs. When implemented properly, cumulative voting and constrained cumulative voting tools can provide a clear and transparent mechanism, showcasing how votes are distributed among candidates or options (Cagliero, 2021).

In the case of Italy, where the cumulative voting method was used, the results of the prioritisation exercise were made available in a document produced by the Ministry of Agriculture (Mazzocchi et al., 2021). The identification of the needs to be included in the CSP was based on 2 central points: robust methodology and a high level of participation. The publication of all the documentation and presentations made during the diagnostic steps was another important aspect to reinforce the overall transparency of the process (Cagliero, 2021).

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?

Voting and prioritisation tools have been mostly used for needs assessment tasks and related activities. However, it could be potentially employed for a number of other tasks and can address all CAP objectives.

These tools are used most often for the need assessment. They are applied for facilitation of a collective discussion and putting in order of priority the identified needs (identified in the former steps) (Cagliero et al., 2021). In the case of Lithuania, monitoring was also mentioned, while Romanian stakeholders mentioned also ex-ante evaluation. In the case of Italy, voting was applied as well for SWOT analysis and the potential of the tool to check consistency between priority levels and the actual allocation of resources was raised (Cagliero et al., 2022). Additionally, the prioritisation process carried out in Italy had the important role of stimulating initial debate on the contents that should be incorporated into the CSP through the various forms of intervention. Besides, voting tools are commonly used in participatory processes such as stakeholder consultations. This method is particularly suitable for collective and participatory decision-making processes, despite some limitations and the potential for biased results. This highlights the multifaceted utility of voting and prioritisation tools in policymaking (Cagliero et al., 2021). Meanwhile according to the Deloitte/AIS report in SI, the tool was conceptually used at a specific stage of the policy process, so this is the only stage that it is applicable at. However, the methodology can theoretically be applied at any spatial/governance scale.

Voting and prioritisation tools were used to contribute to all objectives of the CAP. Generally, the tools can be tailored or re-adjusted to different CAP objectives but their modifications for different policymaking steps are limited. This limitation was also confirmed by the Italian focus group regarding the constrained cumulative voting tool. Among the disadvantages of the cumulative voting tool is the fact that it cannot be modified to add new options once the process has started (Cagliero et al., 2021). As reported during the focus group conducted in Lithuania, the cumulative voting tool allows for obtaining results similar to those of the focus group approach and a clear representation of the stakeholders' perspective, but it could not support other policies. Analysed voting tools are applicable to all administrative level planning, however additional adjustment of the procedures is needed to combine results.

Experience

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

Evaluation Question: How experienced are MS 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?

Currently, only five MSs (Germany, Italy, Lithuania, Romania, and Slovenia) have used this sub-category of tools in the preparation of their CSPs. The MSs in question did not mention the use of such sub-category tools in other policies but the used tools were prepared and tailored to the scope of needs identified within the SWOTs prepared for the CSPs. For Italy, during the focus group, it was specified that constrained cumulative voting approach not only was applied during the previous programming period, but it has been adopted by CREA as part of one of its research projects on organic farming.

Slovenian participants to our survey and the focus group categorised the needs prioritisation tool as one belonging to expert judgment-based approaches. This categorisation stated that the use of such a tool was a response to previous critiques of agricultural policy in the country, which had an extensive objective scope. During the focus group, it was also confirmed that the Deloitte/AIS report tool was used during the needs prioritisation policy step, making it the only applicable policy stage. The estimation of the aggregate priority distribution of needs in previous Rural Development Programs (RDPs) underlines a consistent level of continuity between the two planning periods in identifying some strategic needs and similarity in the overall shape of prioritisation.

In Poland, the Ministry of Agriculture requested to a national research institute to prepare protocol for prioritisation; however, only part of the proposed procedure was applied in practice. The desk research does not give an answer

to the question of previous use of this sub-category when preparing the rural development programmes (RDP) or planning direct payments systems in the previous CAP.

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? What is the level of detail that these tools can reach?

The analysed sub-category of tools can support analysis and policymaking at different administrative levels. This is proven by the fact that in different countries these tools were used at different levels. In Italy, the constrained cumulative voting was applied at the regional and local level (Cagliero et al., 2021), and also confirmed during the Italian focus group. In Lithuania at the national level, yet with the representatives of regional authorities. In the case of Germany, the federal and land ministries participated, thus the national and regional perspectives were included. Similarly, in Spain voting procedures were adjusted to take into account regional differences. Furthermore, according to the Slovenian focus group, the used methodology can theoretically be applied at any spatial/governance scale. On the other hand, according to our survey, the ministries of Cyprus indicated that the concerned tool has not been used to capture different governance levels.

The level of detail these tools can reach depends on the design of these tools, that is what level of detail they envisage. The approach to the scoring range and the choice of experts to be involved in the process of voting leading to prioritisation are the key determinants of the details reached by these tools.

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?

Cumulative voting approaches can be easily used by a wide range of professionals. According to some studies, it was identified as being one of the fastest and most accurate methods that study participants also rated as easy to use (Cagliero, 2021). However, voters and stakeholders need adequate skills and understanding of the system to utilise it effectively, as reported during the Italian focus group, where it was also mentioned the complexity and the level of detail of the tool. Sahni (2007) also conducted a controlled experiment on cumulative voting and found that ease of use and learning could be potential challenges.

The tools that are part of this subcategory are generally easy to use and do not require any prior training or experience, as indicated by our survey respondents. They are often straightforward and easy to implement, making them accessible to a wide range of users with varying levels of expertise or experience in the prioritisation process. However, these tools' rigid structure may hinder flexibility in accommodating evolving needs or unexpected changes, limiting its adaptability in dynamic situations.

Some training might still be necessary for some methodologies the time needed for the training is not significant. It has been implemented by managing authorities mostly and not by an entity specialised in the methodology, which indicates that the tool can be handled by any type of professional. However, the design of the methodology was sometimes made by a different entity than the one implementing it, indicating that building the methodology can require external skills, such as knowledge of facilitation and prioritisation techniques. This was the case for the needs prioritisation approach implemented in Slovenia, which was created by a consulting company and considered as

relatively simple by the participants of the Slovenian focus group, or Poland and Italy, where a research institute was involved in the design.

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 tools that are part of this subcategory are generally available for free and to all users, with no legal or patent-related limitations. However, building a methodology adapted to the decision-making context can be made by an external provider, which implies financial costs for the design of the methodology. This was the case of the needs prioritisation approach implemented in Slovenia, which was created by a consulting company and the Agricultural Institute of Slovenia, and for the cumulative voting approach developed by the Lithuanian Centre for Social Sciences and the Institute of Economics and Rural Development. Factors such as language, ease of use, and availability of support, appear not to limit the availability of the tool. During the Italian focus group, it has been specified that the information related to the use of the tools were uploaded on the National Rural Network website, allowing systematisation and dissemination of all the documentation used during the programming process.

The main limitation for the use of the tools of the subcategory probably has to do with the required financial and time resource associated with the design of a relevant methodology for needs prioritisation for a given context (i.e. issue addressed in the CSP design process). The level of required resources is proportional to the complexity of the methodology. When implementing cumulative voting, voters and stakeholders need adequate education and understanding of the system to utilise it effectively, which might require additional resources for outreach and education. Therefore, tools such as the needs prioritisation approach developed in Slovenia appear as more complex to implement. Meanwhile, the method used for cumulative voting is quite simple and requires less resources than the needs prioritisation approach.

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 sub-category of tools (i.e time necessary to organise and prepare the implementation)? And how much time does it take to implement the process of the tool (i.e. time necessary to implement the tool until the achievement of the intended outputs)?

The efficiency of the tools from this sub-category depends on the details of their design. The stakeholders in the countries which applied them mentioned the fact that the tools were tailored to the resources available including time. A good example is the Polish case where the more complicated method based on hierarchical prioritisation was not used due to lack of time and methodological constrains.

The time needed for the preparation of the implementation depends on the complexity of the tool, its scope and the number of stakeholders needed to be involved. The more people are needed the longer the time necessary to choose the suitable representatives and ensure their participation. For example, according to our survey, the Cyprus Ministry of Agriculture needed up to 1 month to set up and implement a voting and prioritisation tool. According to evidence from the Italian focus group, the cumulative voting approach requires a significant amount of time for preparation and the implementation of the tool's functionalities. Conversely, during the Slovenian focus group, it was specified that the Deloitte/AIS approach generally requires limited time for the implementation process. The main

issues related to this tool are linked to the timing of the preparation of the report, which preceded the adoption of the final CAP legislative acts, and so had to be supplemented with additional work after the publication, resulting in a rather different categorisation and consequently prioritisation of needs.

The time required for the implementation of the tool depends on all the above-mentioned issues related to the time needed for the preparation phase. Moreover, implementation times depends on the number of intermediate steps. The more iterations planned; the more time consuming is the tool.

Cost-effectiveness

Cost-effectiveness refers to the extent to which the tool is able to achieve its outcomes within a reasonable amount of 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 budget can be necessary for the maintenance of the tool or subscription for patent-based tools every year?

The costs of these tools vary significantly depending on the level of their complexity. The first element of the costs is the design of the tool, depending on if the tool is commissioned to an external entity or is design internally by the ministry and/or the institutions supervised by them the cost can vary. The second element is the number of steps that are required to reach the final hierarchy of needs. Besides, there is an issue of the number and diversity of experts who need to be involved and whether their involvement will involve traveling and catering costs. The costs also depend on the scope of the task to be fulfilled, if it encompasses all the CAP priorities it is much higher than when it is limited to a single priority. Therefore, it is difficult to estimate the needed budget and compare the cost incurred by different MSs.

For example, according to the Italian focus group, the implementation of cumulative voting in Italy incurred costs exceeding 2,000 Euro, whereas the partnership approach cost less than 1,000 Euro, as mentioned during the discussion. Similarly, during the focus group conducted in Slovenia, was mentioned that the set-up of the Deloitte/AIS report needs limited financial resources. In Cyprus, our survey indicated that the ministries needed less than 1,000 Euro to implement the voting tool.

Human Resources Efficiency

Human resource efficiency refers to the extent to which the tool is able to achieve its outcomes involving a reasonable amount 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?

The number of people necessary for the implementation of the tool and the analysis of results depends on the scope of the use of the tool and the complexity of it. This sub-category of tools requires multiple persons from different backgrounds and perspectives to be engaged in problem-solving and needs prioritisation (Cagliero et al., 2021). One can expect significant increase of human resources use in case of regionalisation of the CSPs. The key challenge to efficiently use human resources is related to clarity of introduction of the method and to conduct voting in a manner to avoid biases from the effects of social pressure and conformity (Cagliero et al., 2021).

Moreover, the tools from this sub-category require a considerable level of skills of the people involved in the tool design, tough somehow less in the implementation and analysis of the results. In the case of the use of the constrained cumulative voting, according to the Italian focus group there were more than 5 people needed. Similarly, according to our survey, the Cypriot Ministry needed between 5 and 10 people to implement a similar tool.

3.2.3. Multicriteria analysis tools

Summary

***Accuracy**: moderate to high, as the accuracy of these tools relies on precise input data and robust validation processes, despite some limitations in data error prevention and geographical coverage.*

There are some measures in place to avoid mistakes and redundancy in the data. However, the results of these tools are based on well-founded evidence, and the information sources employed are clearly identified and documented. Multi-criteria analysis (MCA) allows decision-makers to consider multiple criteria and their interrelationships, hence accuracy depends on the precision of input data, the appropriateness of criteria selection, and the validity of the scoring and weighting methods employed. The respondents to the survey reported that data validation processes are in place, including quality assessment, cross-verification, and stakeholder input, along with continuous monitoring of data sources. Transparency is also ensured through open access, clear reporting, and reproducibility. None of the respondents indicated that the tools are used to capture information from different governance levels.

***Reliability:** moderate, mainly due to difficulties in replicating MCA results and the quality of stakeholder inputs.*

MCA received an above-average score for reliability, defined as the extent to which a tool's results are consistent across studies addressing the same question. Some of the respondents acknowledged the challenges in replicating MCA results due to the use of various MCA methods, differing stakeholder preferences and inputs, and modelling assumptions that could be implemented across the different studies. These factors may lead to divergent outcomes when addressing the same question, resulting from distinct approaches in handling criteria, weights, and scoring. Regarding the transparency of results, respondents are unsure whether detailed documentation concerning the underlying methodology of the tool is available and accessible to all stakeholders.

***Applicability:** moderate, as these tools are applicable to all policy objectives, but their application to other policies or different governance levels is uncertain.*

According to the results of the survey, MCA tools can contribute to all CAP objectives. MCA exhibits potential for tailoring to multiple objectives and addressing various policymaking steps, as ex-ante analysis, financial allocation, and needs prioritisation, enabling decision-makers to explicitly consider trade-offs and adapt to shifts in priorities or different stages of the policymaking process. On the other hand, respondents state that these tools can be used for all policy steps and have a history of use in previous programming periods. However, respondents also express uncertainty about whether the tools have been applied to other policies. Additionally, the tools are not used to gather information from different governance levels, suggesting that appropriate geographical coverage is not necessarily guaranteed.

***Accessibility:** moderate, due to practice and supplemental resources needed to build the required skills, with extra training for advanced features ranging from 1 to 20 hours.*

The survey indicates that clear information regarding the availability and accessibility of detailed documentation on the underlying methodology of the tool is lacking. This uncertainty also applies to the tool's availability. Various tools have differing levels of complexity, and the amount of training required can vary significantly. Proficiency with these tools is typically achieved through practice and hands-on experience, and is further facilitated by user manuals, online tutorials, and documentation, commonly provided by the developers. For certain specialised applications or advanced features of MCA tools, additional training or expertise may be necessary. According to the respondents, the time required for this additional training can vary, ranging from 1 to 20 hours.

***Efficiency:** moderate, as setting up these tools can take between a month and six months respectively, with a required staff of up to 50 people in certain cases.*

The setup and implementation time for the proposed tools can be lengthy, with setup requiring more than a month. The time required for applying MCA in agricultural policy contexts varies, influenced by factors such as decision problem complexity, data availability, stakeholder engagement, and the decision-making context. Over six months

can be needed. Overall, these tools can involve up to 25 to 50 people. Specific information regarding the costs of implementing MCA is unavailable. However, it is possible to say that the cost for MCA can vary due to factors like using open-source alternatives, which may lower the costs, or the purchase of licenses. The project's complexity, the need for effective management and coordination, the level of expertise and stakeholders involved, as well as the costs associated with communication, reporting, and documentation also influence the overall cost.

Analysis by criteria

ACCURACY

Accuracy analysis considers how successful is the tool 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: 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? Is it possible to easily trace back the information used for supporting the policy choices to its original source? 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)?

In multicriteria analysis (MCA) the process begins with identifying and defining decision-relevant criteria, a subjective step where decision-makers assign weights to reflect each criterion's importance. Numerical values or qualitative scores are assigned to represent the performance of each alternative with respect to each criterion. These scores may be based on available data, expert opinions, or stakeholder input. Sensitivity analysis is often included in MCA to examine how changes in criteria weights or scores affect the ranking of alternatives. This helps identify the robustness of the results and the potential impact of uncertainties in the assessment (Kamali, F.P. et al., 2017).

The integrity of MCA depends on using accurate, validated data to avoid distorted conclusions. It is crucial to ensure data sources are current, relevant, and reflect real-world scenarios. Data traceability is achieved through thorough documentation of data sources, collection, and preprocessing methods, along with metadata and data versioning, to maintain clarity and accuracy from origin to use. (Kamali, F.P. et al., 2017).

Achieving data traceability involves integrated practices to ensure data clarity and accuracy from its origin to its application. This process includes detailed documentation of data sources, methods, preprocessing steps, and metadata to describe data characteristics, with data versioning critical for tracking changes over time. Rigorous data validation, quality assessments, and cross-verification with independent sources or through stakeholder input enhance data accuracy and reliability. These efforts ensure a transparent and credible path for data use in decision-making, reinforcing data's credibility. Transparency, security, and continuous monitoring of data are vital for MCA's reliability, supported by open data access, transparent reporting, and adherence to data protection regulations, fostering reproducibility and compliance, especially for sensitive information.

Regular updates and monitoring of data sources maintain the analysis's relevance and traceability. Robust data management and transparency about data uncertainties and limitations, along with quality reviews and sensitivity analyses, bolster MCA's robustness. This comprehensive approach builds stakeholder trust and supports decision-makers in developing informed, defensible policies and strategies (Kamali, F.P. et al, 2017).

According to our survey, the overall reliability and traceability of the input sources used in the multicriteria analysis received an average score of 4/5. The survey respondents indicated that for MCA there were some measures in place to prevent mistakes and redundancies, such as verification processes, standardised data collection protocols, and data quality audits. Nevertheless, according to the evidence from the Lithuanian focus group, the other used tools such as comparative analysis, expert judgement and socio-economic analysis are accurate and provide valid results. Comparative analysis is carried on a regular basis, also the National Paying Agency provides data on various indicators and after that various policies can be and are updated to reach the best possible outcomes.

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: Are the tools able to capture the diverse facets, variations, or elements within the system or dataset it seeks to represent? Does this sub-category of tools facilitate the involvement of all relevant stakeholder groups while ensuring geographical and sectoral representativeness?

The representatives of MCA tools rely in their ability to accurately represent a system or phenomenon through adequate accuracy and coverage. While MCA is beneficial for supporting policy decisions by allowing the consideration of multiple criteria and their interrelations, its success depends on precise data, suitable criteria selection, and valid scoring and weighting methods. However, MCA has limitations, including potential biases from inaccurate or incomplete data, underscoring the need for rigorous data validation and verification to ensure reliable outcomes. (Abel, E. et al., 2018). Additionally, MCA results is affected by assumptions about criteria relationships, stakeholder preferences, and decision problem behaviour. Sensitivity analyses improve understanding of the decision landscape by evaluating how changes in criteria weights or assumptions influence outcomes.

Comprehensiveness, or coverage, is another essential consideration in MCA. The scope of criteria selected plays a pivotal role in the tool's ability to provide a comprehensive analysis (Abel, E. et al., 2018). Achieving comprehensiveness in MCA involves addressing the challenge of identifying all relevant criteria, considering spatial and temporal factors for problems across various regions or periods. Engaging stakeholders enhances this by bringing diverse perspectives for a more holistic view. Methods like the Analytic Network Process (ANP) further this goal by examining interactions between criteria, thus providing a nuanced understanding of the decision context. Recognising trade-offs and synergies among criteria also enriches the analysis, aiding decision-makers in making more informed choices. None of the respondents declared that the tools are used to capture information from different governance levels. This suggests that either the geographical coverage is not ensured or that it wasn't needed.

In conclusion, achieving an accurate and comprehensive representation of a system, dataset, or phenomenon for policy choice decisions through MCA requires careful consideration of factors such as data quality, modelling assumptions, criteria selection, and stakeholder involvement. Thorough validation, sensitivity analysis, and a collaborative approach with stakeholders contribute to a more accurate and comprehensive depiction of the decision context. The degree of representativeness achieved is influenced by the commitment to inclusivity and the level of stakeholder involvement throughout the decision-making process.

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?

Replicability, defined as the extent to which a tool's results are consistent across studies aimed at answering the same question, is a crucial consideration in MCA. The consistency of MCA study results can be influenced by several factors, and expectations for replication may vary based on the decision problem's nature, the chosen MCA method (Cinelli, M et al., 2020), and the specific context of the studies.

Considerations for replicability in MCA include methodological replication, where applying the same MCA method to the same decision problem with identical criteria, weights, and data ideally yields consistent results, assuming the decision problem remains unchanged. Replicating MCA studies using different datasets introduces variability, with

data source, quality, and availability impacting results. Using a different MCA method to address the same question may lead to divergent outcomes due to distinct approaches in handling criteria, weights, and scoring (Dehe, B and Bamford, D., 2015). However, results could be reproduced by repeating the process of implementation of the tool under the same conditions (Gongora-Salazar, P. et al., 2023).

Stakeholder preferences, the dynamic nature of criteria, and external contextual factors like policy or economic changes can lead to variability in MCA outcomes across studies. Differences in modelling assumptions can also affect the consistency of results. Sensitivity analyses are crucial for assessing the robustness of the findings against changes in criteria weights, scoring, or assumptions, highlighting the reliability of the results, since results sensitive to variations may impact expectations for replication (Mardani, A. et al., 2015).

In summary, while achieving exact replication of MCA results across studies may be challenging due to these factors, the goal is to ensure robust, transparent, and sensitive results. MCA practitioners should carefully document methodologies, assumptions, and data sources and conduct sensitivity analyses to understand potential variability.

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?

The respondents to the survey reported that they do not know whether detailed documentation concerning the tool's underlying methodology is available and accessible to all stakeholders. The extent to which information about underlying methodologies, data, and results of a tool is available and accessible constitutes the basis of transparency. Additionally, the openness of the tool to external review is a crucial aspect of studies employing an MCA approach for policy choice decisions (Dyckhoff, H., Souren,R., 2022).

Several observations on transparency and openness to external review in MCA studies can be made, publication standards in academic journals often mandate transparency and reporting. Authors are typically required to furnish detailed descriptions of their methodologies, encompassing criteria selection, weighting methods, and data sources. Peer review processes, involving external experts evaluating research methods and results, are commonly employed in such publications (Dyckhoff, H., Souren, R., 2022; Dean, M., 2022).

Transparent MCA studies typically incorporate a dedicated methodology section outlining the step-by-step analysis process. This section may cover criteria selection, weighting, scoring methods, the chosen MCA tool, and any sensitivity analyses performed. Studies valuing transparency often include information on data sources, collection methods, and preprocessing steps, enabling readers to assess the quality and relevance of the data used in the MCA (Dyckhoff, H., Souren, R., 2022)

Transparent MCA studies often feature sensitivity analyses, exploring the impact of changes in criteria weights, scoring, or assumptions on the results, contributing to the assessment of robustness. Some MCA studies are published as open-access articles, promoting broader access to the research community and the public (Dean, M., 2022). A trend towards data sharing and making datasets available facilitates external scrutiny and validation. Transparent studies explicitly document any assumptions made during the analysis, including those related to stakeholder preferences, model parameters, or simplifications in the decision-making process. Transparent studies acknowledge and discuss the limitations of the MCA approach and the analysis, providing readers with a clear understanding of the study's constraints and potential areas for improvement. Some studies include supplementary materials, such as appendices or online resources, providing additional details on the MCA methodology, raw data, or extended results, thereby enhancing transparency and supporting external review.

Transparent MCA studies align with the push for reproducibility in research, providing the necessary information and code (if applicable) to enable others to replicate the analysis and verify results. While many MCA studies strive for transparency, practices can vary. Authors and researchers play a crucial role in ensuring transparency, and journals or platforms contribute by setting standards for reporting. The willingness to share methodologies and data and to

engage in external review processes enhances the credibility and reliability of MCA studies in policy decision-making.

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?

The concept of multifunctionality within the context of MCA highlights its capacity to simultaneously pursue various objectives and navigate multiple stages of policymaking. According to our survey respondents, these tools contribute to all CAP Objectives. This versatility enables decision-makers to tailor MCA to achieve diverse goals and address different aspects of policy formulation and implementation concurrently. In terms of policy steps, MCA has been used for needs assessment, ex-ante Analysis and budget allocation.

The flexibility to alter underlying assumptions in MCA is influenced by several factors, enhancing its suitability for multifaceted policy decisions. MCA's adaptability stems from its ability to incorporate a wide range of criteria that represent different dimensions of a policy decision, facilitating a thorough evaluation. This adaptability is further augmented by the dynamic modification of criteria, which allows for the inclusion of evolving priorities or new policy considerations. Through MCA, decision-makers can explicitly assess trade-offs between objectives, employing criteria weighting and scoring methods to mirror the relative significance of each goal within the decision-making process.

The process of adjusting criteria weights enables the prioritisation of certain objectives over others, reflecting shifts in policy emphasis or the requirements of various policymaking phases. Moreover, stakeholder engagement plays a crucial role in MCA by introducing diverse perspectives and objectives into the decision-making process. This involvement aids in identifying pertinent criteria and gauging their importance, thus facilitating consensus building and aligning the MCA approach with the collective aims of all stakeholders, thereby bolstering its relevance to a wide array of objectives and policy stages.

MCA methodologies often incorporate sensitivity analysis, which allows decision-makers to evaluate how changes in assumptions influence outcomes, thereby increasing the method's adaptability to different scenarios and considerations (Munda, G. et al., 2020). Scenario analysis further enhances this adaptability, enabling the exploration of various futures by modifying underlying assumptions to better address the multifaceted nature of policymaking.

The integration of MCA with other decision support methods, such as cost-benefit analysis or scenario planning, expands its capability to tackle diverse objectives and stages of policymaking. Its iterative nature supports the continuous refinement of assumptions and criteria, promoting an adaptive and responsive policymaking approach.

In essence, the inherent flexibility of MCA makes it an invaluable tool for pursuing multiple objectives and managing various stages of policymaking (Munda, G. et al., 2020; Dean, M., 2020). Its capacity for adaptation is underpinned by the modification of criteria, adjustment of weights, incorporation of stakeholder preferences, and execution of sensitivity analyses, collectively ensuring its applicability and effectiveness in complex policy environments.

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?

Our survey respondents indicated uncertainty regarding whether the tools have been applied to other policies. They confirmed that MCA was utilised in previous programming periods. Nevertheless, there are several examples of potential applications of the proposed tools.

This is the case of the European Joint Research Centre (JRC) that has developed a software tool, called Socrates (SOcial multi-CRiteria AssessmenT of European policieS)[45], applicable to agriculture and rural development for ex-ante policy impact assessment. It implements the main principles of social multi-criteria evaluation (SMCE) and has three main components: multi-criteria, equity and sensitivity analyses.

About agriculture specific applications, a Lithuanian study[46] devised a system of indicators to assess the impact of CAP direct payments on the socioeconomic viability of small-scale farms. This evaluation involved conducting a survey among experts and implementing a multi-criteria analysis using the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) method. This method allows to compare the emission intensity of the agricultural sector across 28 EU MSs (MS), which is a suitable assessment method as well as an appropriate tool to be for the identification of preferred solutions based on multiple criteria (Dace, e., Blumberga, D., 2016). Whithin the study, the approach consolidated indicators into a unified metric for social sustainability of small farms, linking changes in CAP direct payment system to these metrics. Findings indicated that from 2004 to 2016, the Direct Payments system negatively affected local production diversity and failed to support small farm incomes effectively, undermining social sustainability due to market price and income dynamics. However, post-2013 adjustments in CAP Direct Payments, with increased support for small and young farmers, improved social sustainability indicators, suggesting better social conditions for these farms. The study also noted that outcomes were significantly influenced by the weighting schemes used, reflecting the subjective nature of expert opinions.

Finally, it is also available a multicriteria model to study the social impacts of the CAP reform (“decoupling”) and Water Framework Directive in Greece (Manos B., et al., 2011). The model estimates the farmers’ utility function taking into account various conflicting criteria that can explain farmers’ behaviour (e.g. maximisation of farm income, risk minimisation, minimisation of labour, etc.). The model is further used to simulate the impacts on social sustainability by estimating a selection of social indicators.

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? What is the level of detail that these tools can reach?

Granularity, within the context of MCA, denotes the tool's ability to deliver results across a spectrum of spatial resolutions, ranging from local to regional, national, and even EU levels. Although none of the survey respondents declared that the tools are used to capture information from different governance levels, suggesting a possible gap in geographical coverage or a lack of need, it is demonstrated that MCA tools are particularly effective for addressing

[45] Further information on the tool can be found at the following link, where reference is made to a number of applications outside of agriculture: https://knowledge4policy.ec.europa.eu/modelling/topic/social-multi-criteria-evaluation-policy-options_en/socrates_en

[46] https://www.researchgate.net/publication/332323470_Who_Benefits_from_CAP_The_Way_the_Direct_Payments_System_Impacts_Socioeconomic_Sustainability_of_Small_Farms

agricultural policy issues that demand analysis over various spatial scales, notably at regional and national levels (Velasquez, M., Hester, P. T., 2013). Its relevance is due to its ability to accommodate diverse criteria across different spatial levels, such as climate, soil types, and socio-economic conditions, which vary by region or nation. It allows decision-makers to assign weights to these criteria based on their importance at various levels, ensuring that decisions reflect regional or national priorities and address unique local challenges. Stakeholder participation enriches MCA by integrating expertise, preferences, and local knowledge, enhancing the contextual relevance of policy decisions. MCA facilitates the evaluation of policy alternatives tailored to the specific needs and characteristics of individual geographic areas.

Some MCA tools, such as the Analytic Network Process (ANP), are specifically designed to manage interactions and dependencies among criteria and alternatives across spatial levels, crucial for complex agricultural policy issues that span regional and national dynamics. MCA facilitates sensitivity analysis, enabling decision-makers to gauge how variations in criteria weights or scores influence policy rankings in different spatial contexts, thus assessing policy recommendation robustness. Integrating MCA with geospatial data, such as GIS, improves the analysis of spatial criteria, enhancing understanding of spatial influences on policy decisions. MCA supports strategic planning by evaluating policy trade-offs and synergies across levels, aiding in the formulation of coordinated policies that address spatial variations and achieve national goals.

In conclusion, MCA is exceptionally well-suited for analysing agricultural policy issues across varied spatial levels. Through the customisation of criteria, involvement of stakeholders, and sensitivity analysis across regions and nations, MCA presents a flexible and comprehensive approach to decision-making within the intricate and diverse landscape of agricultural policy.

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?

The training required to effectively utilise MCA tools depends on the tool's complexity, the user's understanding of decision analysis principles, and the application context. Users already familiar with concepts like criteria identification, weighting, scoring, and aggregation are likely to comprehend MCA tool functionalities more easily. Since different MCA tools exhibit varying levels of complexity and users' familiarity with decision-making processes and criteria identification provides a solid foundation. According to our survey respondents, MCA can require previous experience/training. However, this is limited to 1-20h. Some, like Simple Additive Weighting (SAW), are relatively straightforward and may require minimal training. Others, such as Analytic Hierarchy Process (AHP) or Analytic Network Process (ANP), can be more intricate and may necessitate a deeper understanding of mathematical concepts and pairwise comparisons (Velasquez, M., Hester, P. T., 2013).

The training needed for MCA tools varies with the software's user interface, functionalities, and the user's familiarity with its methodology, including criteria identification, weighting, aggregation, and decision-making processes. User-friendly tools may be easier to learn, whereas others require deeper understanding and practice. Effective use of MCA, especially involving stakeholder participation, may also necessitate training in facilitation skills and stakeholder engagement methods. Gaining proficiency typically involves hands-on application to real-world scenarios, understanding the tool's capabilities and limitations. Specialised applications in fields like healthcare or environmental management might need additional, domain-specific training. Many MCA tools offer support like manuals, tutorials, and documentation, with training programs and workshops available for comprehensive learning.

In summary, the amount of training required to learn MCA tools can vary widely. While individuals with a background in decision analysis may quickly grasp concepts and applications for simpler tools, more complex tools or those with advanced features may necessitate a more substantial investment in training and practice. Continuous learning and practical application are essential for users to become proficient in effectively utilising MCA tools for decision-making in diverse contexts.

According to the focus group conducted in Lithuania, other multicriteria analysis tools, including comparative analysis, expert judgement or socio-economic analysis are easier to use as there are stricter methodological requirements to encompass.

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?

According to our survey, respondents do not know whether the tools themselves and detailed documentation concerning the underlying methodology are available and accessible to all stakeholders. From the Lithuanian focus group emerged that tools such as comparative analysis, expert judgement or socio-economic analysis have stricter methodological requirements to encompass and are based on expert knowledge.

The accessibility of MCA tools can vary depending on several factors. Some MCA software tools are open-source and freely available for download. These tools often provide a wide range of functionalities for conducting MCA, such as implementing different decision-making methods, sensitivity analysis, and visualisation capabilities. Many organisations, academic institutions, and government agencies provide publicly available guidance documents, manuals, and technical reports on MCA methods and applications. In an EU context, some useful general guidance on appropriate methodologies can be found in the Better Regulation Toolbox[47], as well as resources related to the SOCRATES model[48].

Overall, while there are numerous freely available resources for learning about and applying MCA techniques (Manos, B., et al., 2011; Munda, G., 2021; Munda, G., 2004; Munda, G., 2008; Siciliano, G., 2009; Volkov, A., 2019), access to more advanced tools and support may require investment or collaboration with specialised organisations or software providers. The complexity of MCA methods and techniques may pose challenges for users without a strong background in mathematics, statistics, or decision analysis. Availability of resources in languages other than English may limit accessibility for non-English speaking users, particularly in regions.

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.

[47]https://commission.europa.eu/law/law-making-process/planning-and-proposing-law/better-regulation/better-regulation-guidelines-and-toolbox_en

[48]https://knowledge4policy.ec.europa.eu/modelling/topic/social-multi-criteria-evaluation-policy-options_en/socrates_en

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

Time efficiency in utilising MCA involves achieving desired outcomes within a reasonable period, including both setup and implementation phases. According to our survey, the time to set up the tool exceed a month, in some cases requiring over six months for implementation. The duration to apply MCA in decision-making varies widely, depending on factors such as the complexity of the decision issue, choice of MCA tool, data availability, stakeholder involvement, and users' familiarity with MCA methodologies. Simpler tools like Simple Additive Weighting (SAW) allow for quicker application, whereas more complex tools like Analytic Hierarchy Process (AHP) or Analytic Network Process (ANP) require more time due to their more complex mathematical and comparative analyses (Velasquez, M., Hester, P.T., 2013).

Data availability significantly influences the MCA process timeline. If decision-makers have all necessary data readily available, MCA can be executed efficiently. However, substantial data gaps or the need for extensive data collection and processing can delay the process. Stakeholder engagement, through consultations and workshops, is vital for adding diverse perspectives and ensuring the decision-making process's relevance and acceptability, but it also requires careful planning, potentially extending the MCA timeline. The process efficiency also depends on the decision-makers' and analysts' familiarity with the MCA tool. Experienced users can proceed more quickly, whereas newcomers may need extra time for training. The urgency of decision-making and computational resource availability are further considerations; urgent decisions may call for a streamlined approach, while more extended timelines can accommodate detailed analyses. Complex tools needing significant computing power necessitate adequate resources for efficient data processing. Additionally, MCA may involve iterative steps like refining criteria or adjusting weights, crucial for maintaining analysis accuracy and relevance but possibly prolonging the process. These iterations, especially important in adapting the analysis for specific agricultural policy decisions across different regions or times, can extend the MCA duration.

In summary, the time required for applying MCA in agricultural policy contexts is context-dependent. Factors such as decision problem complexity, MCA tool characteristics, data availability, stakeholder engagement, and the decision-making context collectively influence the duration of the analysis. A well-planned and transparent process, along with effective communication, can help manage expectations regarding the time commitment associated with applying MCA in agricultural policy contexts.

Cost-effectiveness

Cost-effectiveness refers to the extent to which the tool is able to achieve its outcomes within a reasonable amount of 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 budget can be necessary for the maintenance of the tool or subscription for patent-based tools every year?

Cost-effectiveness, in the context of applying the MCA for agricultural policy decisions, relates to achieving desired outcomes within a set and reasonable budget. Costs can fluctuate based on the complexity of the decision problem, the selected MCA tool, data availability, the level of stakeholder involvement, and the expertise required. The exact costs for implementing these tools can vary and is unknown to the respondents of our survey.

Some MCA tools may entail licensing fees or software costs, impacting the overall budget. Data collection and processing costs are significant, with the availability of comprehensive data potentially lowering costs, while extensive collection efforts may increase them. Licensing fees or software costs of MCA tools, along with the level of expertise and familiarity required for effective application, can impact the overall budget. Stakeholder engagement methods, which includes aspects such as number and diversity of stakeholders, engagement methods, and the need for consultations or workshops, but the necessity for expert consultations, which includes can also affect costs, though they improve analysis quality.

Iterative processes for refining criteria and stakeholder feedback, the complexity of decision problems, and the computational resources needed for analysis contribute to costs. The choice between open-source and proprietary

MCA software affects costs as well since the computational resources needed for data processing and analysis contribute to costs. Tools with complex mathematical models may require sufficient computing power, incurring additional expenses. Customising analyses for specific regions or times and expenses for communication, reporting, and expert consultations add to the budget. Costs associated with communication, reporting, and documentation should be considered. Transparent communication of the MCA process and results may require resources for reports, presentations, and other materials.

Despite these expenses, the investment in MCA can enhance decision-making and contribute to developing effective agricultural policies, with its cost-effectiveness measured by the added value and improved policy outcomes.

Human Resources Efficiency

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

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

Overall, multicriteria analysis tools, including comparative analysis, expert judgment, and socio-economic analysis, have stricter methodological requirements and thus require expert knowledge and individuals with appropriate skills. Human resources efficiency in applying MCA for agricultural policy decisions relates to achieving outcomes with an optimal number of personnel. The necessary team size can fluctuate due to the decision problem's complexity, the MCA tool used, stakeholder involvement, and resources. Our survey indicates that MCA typically needs fewer than 5 people, but in some cases might require up to 25 and 50 people.

Complex issues might demand larger, multidisciplinary teams. The choice and complexity of the MCA tool also affect team size: simpler tools may need fewer people, while more complex ones require specialised expertise. Data collection and processing needs can dictate whether a larger team is necessary, especially for extensive or specialised data handling. Stakeholder involvement can expand the team size needed to ensure comprehensive communication and input integration. The team's expertise with MCA methods plays a significant role: experienced members can reduce the need for a larger team, whereas specialised tool requirements might necessitate additional expertise.

Iterative processes and the need for clear communication throughout the MCA process could influence the preference for a larger team. The decision timeline and available resources might also necessitate a bigger team to meet tight deadlines or manage extensive analyses. Furthermore, decisions covering various spatial levels may require team members knowledgeable in regional specifics. The diversity needed for stakeholder engagement and addressing the agricultural policy's multifaceted nature may call for a team capable of handling different policy aspects efficiently.

3.2.4. Key findings

In this section, we highlight the key aspects emerging from the policy choices supporting tools. We conclude by identifying the strengths and weaknesses of this category of tools.

Policy choices supporting tools, which include Voting and Prioritisation (VP) tools and Multicriteria Analysis (MCA) tools, aid in managing intricate systems characterised by interdependencies, trade-offs, and interactions among components, especially in scenarios with various policy alternatives. Information is obtained from diverse sources involving multiple decision-makers. These tools are key in reaching a unified, cohesive, and logically sound policy choice.

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

Accuracy is assessed as limited-to-moderate (VP) and moderate-to-high (MCA). More precisely, accuracy in voting tools is limited by oversimplification and strategic behaviours. Data-driven approaches can enhance accuracy but can face challenges, which may be related to limitations in data quality, subjective interpretation and incomplete information. When it comes to MCA, the tools rely on more precise data with moderate to high accuracy, emphasising transparency and validation processes. However, capturing information from different governance levels is an unaddressed aspect, potentially affecting geographical coverage.

Reliability is judged as limited-to-moderate (VP) or moderate (MCA) for this category of tools. Voting tools are assessed between limited to moderate due to their structural complexity and subjectivity, which limit the replicability of results. On the other hand, MCA tools are deemed moderately reliable as their results vary across methods, stakeholder inputs, and assumptions, impacting consistency. Additionally, transparency in MCA methodology documentation is uncertain, affecting result reliability.

Applicability is assessed as moderate-to-high (VP) or moderate (MCA). These tools demonstrate potential for tailoring multiple CAP objectives and various policymaking steps, and few MSs already use these tools. Nevertheless, MCA applicability to different governance levels remains uncertain.

Accessibility is assessed as moderate-to-high (VP) or moderate (MCA). Voting tools face constraints in introducing new options, with only five MSs using it, but it does not require advanced skills. MCA requires practice and supplemental resources to build the necessary skills, with extra training for advanced features.

Efficiency is assessed as moderate-to-high (VP) or moderate (MCA). For voting tools, efficiency is usually moderate to high, depending on design and resource tailoring. Costs are influenced by complexity and scope, requiring skilled personnel. On the other hand, MCA tools exhibit moderate efficiency, with setup times exceeding a month and six months, respectively. MCA implementation is influenced by factors like complexity, generally requiring few people, although in certain cases they can involve up to 50 persons. Variable costs in MCA depend on factors like project complexity and stakeholder involvement.

Policy choices supporting tools are crucial in achieving a coherent consensus for a well-reasoned policy decision but they also have some notable challenges. These tools have limited to moderate accuracy, reliability, applicability, and accessibility, mainly due to over-simplification, complexity, and uncertainty in application to other policies. On the other hand, this tool's efficiency is judged as moderate to high depending on factors such as resource tailoring and complexity. Still, these tools can be beneficial when dealing with complex decisions involving multiple options and views, such as needs assessment, intervention setting or financial allocations. By nature, these tools can involve multiple actors and can be complemented with stakeholder needs assessment tools. Using policy analysis tools can also help improve the accuracy and reliability of results provided by policy choice supporting tools.

3.3. Policy analysis tools for evidence-based decisions

3.3.1. Inventory of policy analysis tools for evidence-based decisions

Table 7 below shows examples of tools belonging to the different sub-categories, according to our inventory of tools. The full list of tools is available in the Tools4CAP online inventory.

Table 7. Sub-categories of tools and examples of tools from the Tools4CAP inventory.

Sub-category of tools	Example of tools identified
Statistical methods	• Statistical and environmental data analysis (Bulgaria) • Statistical Socio-economic Context Analysis (Poland) • IOI matrix (Netherlands)
Simulation models	• SiTFarm typical farm model tool (Slovenia) • Eco-Scheme Farm simulation tool (Netherlands) • CAPRI (Joint Research Centre of the European Commission) • Eco-Scheme Farm Simulation Tool (Netherlands) • Farm Income Calculation Tool (Netherlands)

Policy analysis tools have been employed to support different tasks of the CSP design and monitoring. Table 8 below provides an overview of the number of MSs where a specific sub-category of tools was employed for each task of the CSP design and monitoring. More details are available in the online inventory of tools[49].

Table 8. Use of policy analysis tools across the CSP design and monitoring (no. of MSs in which the tools were employed for the specific task).

	Statistical analyses	Simulation models tools
CSP design tasks		
Socio-economic context analysis	3	
SWOT analysis	3	
Needs assessment	1	
Interventions setting		3
Target setting		1
Financial allocations		1
Ex-ante analysis and SEA		
Stakeholders' consultations		
CSP monitoring tasks		
Performance review	1	
Beneficiaries' compliance		
Evaluation		

The following sections presents the results of the evaluation specific for each sub-category of tools.

[49] https://www.tools4cap.eu/tools/

3.3.2. Statistical analysis tools

Summary

***Accuracy**: moderate, while many data sources can ensure fair accuracy and statistical methods exist to deal with accuracy issues, accuracy remains subject to the level of analysis, type of data source, and sample size.*

Statistical analyses generally rely on reliable data sources, such as the FADN, and can be quite accurate. However, some concerns might arise for the validation of data, which has to do with their generalisation, especially for macro-level analysis (e.g. involving more countries). Statistical analyses can usually be adequately representative, but this is not always the case as they cannot be used for all cases and at all times. Key factors affecting representativeness are the data source (e.g. official statistics versus ad-hoc surveys) and sample size. However, several statistical methods exist to assess and deal with the issue of representativeness.

***Reliability**: moderate to high, due to their replicability and standardisation, fostering consistent results, although different methods may yield varying outcomes.*

Replicability is generally high in statistical methods, particularly those related to descriptive statistics, mainly due to the fact that statistical methods are standardised to enhance comparability. The advantage of standardised statistical methods for data analysis is that whenever those tools are applied to the same data population they lead to the same results. However, when applying different statistical techniques, results might differ. Statistical analyses can be made readily transparent, although in some cases, statistical software might be a black box, and the statistical techniques used are not always clear. To a significant extent, the reliability of statistical analyses depends on the quality of the data sources used.

***Applicability**: moderate to high, they can be applied for different purposes and MSs own broad experience with their use, whereas data at different granularity are available to apply these tools at different scales.*

While statistical analyses can potentially address all policy objectives, their application remains limited primarily to purely analytical tasks, such as socio-economic context and SWOT analyses. Yet, they may be used to support and verify other decisional tasks, such as interventions setting. MSs have broad experience with statistical tools, although this is often limited to the socio-economic and SWOT analysis stages of the CSP design. The granularity of statistical analysis largely depends on the available data, which can be hindered in more aggregated official statistics, but farm-level and yearly data sources are also available.

***Accessibility**: moderate to high, while statistical tools are made available through several software, skills and training may be required for adopting them and interpreting the results.*

Many statistical analyses can be easily implemented in practice without advanced experience or training. Existing statistical software products, as well as several technical guidelines, make implementation accessible. Yet, more advanced skills and training can be required to select the tool and interpret the results, mainly due to the underlying theoretical assumptions that need to be considered, which can make these tools moderately accessible. Except for most advanced statistical models, statistical techniques are generally available through several (sometimes open-access) statistical software along with abundant technical documentation.

***Efficiency**: moderate to high, as there are less costly, time-demanding and require less professionals compared to other simulation models.*

In general, due to the range of available data sources, statistical software and ease of use of many methods, statistical analyses are time-efficient. Instead, significant time would be required for setting up more advanced analyses and for ad-hoc data collection. While there might be costs related to licenses or training, compared to other simulation models, statistical analyses are overall quite cost-effective. Generally, compared to other tools, fewer professionals and training time are required to perform statistical analysis.

Analysis by criteria

ACCURACY

Accuracy analysis considers how successful is the tool 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 back the information used for supporting the policy choices 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?

The accuracy of statistical analyses depends largely on the reliability of the data covered and selected for this purpose. Overall, statistical methods rely on data representing real fact, as mentioned during the Latvian focus group, with some data sources not providing a complete picture of some categories farms (e.g. small farms). The most common statistical methods that can be used at the micro level are from the so-called descriptive statistics, which can be easily complemented with some of the most frequently used statistical tools for data analysis. They are the most widespread, easy to apply and accessible to specialists and experts implemented in a wide range of analyses. Each of the methods of descriptive statistics has very well-described advantages and setbacks that deal with the accuracy of the results and the validation of the yielded outcomes (Koop, 2013).

Descriptive statistics is generally sufficiently reliable and accurate, and the features are predominantly related to the preparation for work with the corresponding methods. This has also been confirmed by the conducted survey, where the responded also reported that there are measures in place to avoid mistakes and redundancy in the input data and that the obtained data by the tools are validated. Econometrics is often employed at farm level, such as the case of behavioural econometrics which study the factors leading so certain farmers' behaviours (Koop, 2013).

Very often, when working with the data at the farm level, the most commonly used and accessible database is the Farm Analysis Data Network (FADN), which includes both the individual records of the farms covered in the observation and the aggregated results after extrapolation and weighting of the farm sample. However, statistical analysis can be performed using other national statistics or ad-hoc surveys (European Court of Auditors, 2023).

The validation of the results obtained from the statistical analyses is a sensitive issue, which depends on generalisation. Along with that, due to statistical consolidation, which is a procedure undertaken in statistical analysis, can also rise discontent that data is either incorrect or that it is not extrapolated properly (OECD, 2008). During the Dutch focus group, for example, it has been mentioned that the results obtained from the IOI matrix tool are further validated by researcher and policy makers.

At macro level (i.e. multiple countries) the accuracy issue has other implications. With respect to econometrically estimated behavioural equations, an accuracy criterion such behavioural equations should satisfy is that they can track the past to a reasonable extent (e.g. the goodness of fit R-square indicator) (Koop, 2013). A strength of econometrically estimated equations, that are based on economic theory (e.g. supply and demand systems), that are secured to satisfy basic behavioural properties (e.g. downward sloping demand curves, or negative own-price responsiveness of demand, and the like), which are imposed on the inference process (e.g. by means of parameter constraints) is that they have the accuracy to reflect normal and expected economic behaviour.

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 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 it seeks to represent?

The representativeness is a central issue in working with quantitative methods. The representativeness per se as an adjective has no meaning if the object and goals to which it refers is not specified and bound in particular. Thus, the correct question should be representative of what. It is also assumed that always when any statistical analysis is undertaken, it depends on the type of sample and observations covered and how the extrapolation and multiplication of the outcomes is done. However, it should be considered that there are underlain methods for measuring the standard error, confidence level, and confidence interval, different categories that complement the issue of representativeness (Gilliland, 2010). Overall, from the survey respondents indicates that data align with the desired result level.

There are adopted standards for statistical significance, especially when working with official databases, such as Eurostat and FADN. The issue of representativeness is more sensitive when different kinds of surveys, field studies, interviews are used, which is very often needed instruments in order to fill up the data gaps. However, it should be envisaged whenever working with a sample that there is a relationship between the sample size and the standard error and hence the confidence interval. The smaller the sample, the larger the standard error is supposed to be and the lower the confidence interval that the sample possesses, and vice versa (Gilliland, 2010).

Statistical tools for descriptive analysis, as well as other data analysis tools, are reliable and representative, but not always universal. This means that they cannot be used for all cases and at all times. Usually, statistical methods by default are not flexible and sensitive because they work with standardisation, which implies some generalisations. It is very important when working with the most widespread methods of statistical data processing that the specialists who interpret the obtained results are well known and trained statistically in order to differentiates the constraints and specifics that might be hidden in statistical outcome extraction.

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?

Replicability is one of the advantages of applying statistical methods, particularly those related to descriptive statistics and data analysis, thanks to the standardisation of these methods, which, while restricting flexibility and applicability, enhances comparability. The advantage of standardised statistical methods for data analysis is that whenever those tools are applied to the same data population they lead to and bring the same results (Killen, 2007). The presence of established methodologies and protocols and their contribution in allowing replicability has been also acknowledged by the participants to our survey. However, when using other methods of analysis, especially specific models or estimation methods, discrepancies in the results may emerge. However, it does not immediately mean that the statistically generated results are not valid or reliable because the statistical results are always standard and are performed under certain conditions characterised by predefined constraints (Koop, 2013).

The methods of descriptive statistics are relatively simple, and it is not usual to obtain different results when using other tools. As for some of the popular and widely used data analysis tools, such as regression, variance, covariance,

trends, etc, it can be noticed that in certain cases different results are obtained even when working with the same database. Therefore, it is considered and recommended that in-depth analysis cannot rely only on the statistical methods, which have their setbacks, and those methods should be complemented with other specific and very often well-adapted quantitative methods for data analysis (Koop, 2013).

If the replication means to repeat the empirical operations and to record data that support the original findings, this can be implemented in the research process by testing different hypotheses that give diverse answers to the same question. From a statistical point of view, methods for checking hypotheses are available and should be considered as a suitable research tool for analysis and conclusion synthesis, especially when dealing with samples (Killen, 2007).

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?

The results obtained by the statistical tools are transparent and clear although sometime raise the question whether the relevant data and dependencies are investigated (Wagenmakers et al., 2021). Whenever statistical tools are applied through software (Excel, R, Stata, etc), it can be difficult to verify if the data is correctly selected and the relationship between indicators are settled properly (Koop, 2013). However, the transparency in using standard statistical tools is quite adequate compared to independent and work specifics models and quantitative. Due to some additional issues of pre-data collection and calibration and patterning, it is recommended to give reference to data source, whereas by exposing the analysis algorithm, analysts provide the basis for their statements and received outcomes (Bleich & Pekkanen, 2015). Ideally, data and algorithms for statistical computation should be shared publicly, freely and in a manner that facilitates reuse. For instance, during the Latvian focus group, it has been mentioned that statistical analysis tools are based on regularly updated data, which is accessible through Rural Support Service and other databases in Latvia.

The transparency of the statistical analysis is also achieved through standardised data collection method because the statistical research covers not only tools for data processing but also tools for collecting data. When researchers share the results from only a single data source, they may unintentionally hide important information. If a result proves sensitive to preprocessing choices, this warrants scepticism and may initiate a debate on the importance and plausibility of relevant data preprocessing choices (Wagenmakers et al., 2021).

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?

Statistical analysis tools are predominantly applied at the stage of socio-economic context and SWOT analysis and ex-anet analysis during the CSP design. Through these tools, it is possible to analyse the situation and the environment that is the subject of the research and policy initiatives (Koop, G., 2013). These tools can be used to

verify results, proposals, and enacted decisions. Thus, when applied, statistical tools are not only an analytical tool but also have room for implementation whenever the relevance and potential effects of programming are sought to be verified. However, in term of policy objectives, statistical analyses can have a wide application. According to our survey, statistical analyses were used in relation to all CAP objectives.

Although statistical methods were not applied in several stages of the CSP design, the wide range of functionalities provides the end user with a greater scope of possibilities to monitor the overall performance of the tools. This includes a rich library of statistical algorithms, text analysis, software extensibility through add-ons, and data integrations. Statistical analysis is a key tool at the investigation stage, and in the evaluation and pre-implementation stage of the policy process, but it also facilitates the task at any time when data must be processed. In this relation, descriptive statistics and statistical data analysis methods are some of the most widespread and useful tools, which, due to different objectives in policy planning, might be adopted to analyse the situation, verify the results, and validate the collected data (OECD, 2008).

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?

All MSs have, to a very large extent, the capacity and experience to implement and work with statistical analyses. When the organisations in charge, especially for the elaboration of the socio-economic context and SWOT analysis, are selected by the authorities, usually one of the criteria and conditions is to have experience and capacity to work with such tools (Brenna, J. et al., 2021). The statistical methods can be applied for different purposes and at various stages of policy elaboration. Always when the work is related to data collection and diagnostic research, the authorised institutions resort to sufficiently experienced and appropriate teams to carry out such a task (Koop, G., 2013).

The level of application of statistical methods in individual MSs differs, and the stages at which they are applied in policy development also differ between MSs. The reasons for diverse application are the approaches particular MSs have, as some MSs rely more on statistical analytical tools, while others more on quantitative models or qualitative tools. Bearing in mind that it is required for all MSs to observe and investigate the context indicators and to proceed with analyses almost every time important policy decisions are undertaken, statistical analyses are carried out, which has created MS experience and competences (Gonzalez-Martinez, A.R. et al., 2021). In general, the survey indicates that some statistical analysis tools have been used in the previous programming period, specifically in Czechia.

Granularity

Granularity refers to the extent to which the tool can provide results at different level of spatial resolution, for example 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 statistical tools are appropriately adopted and used at different levels and aspects of the analytical process. From a statistical point of view, most descriptive statistical methods and those for data analysis are very precise quantitative instruments. Granularity also depends on data availability (Maier, J.F. et al., 2016), often at regional and country levels, as reported during the Latvian focus group the farm level, but also farm level in specific cases. However, this is not the case for the IOI matrix, which is applicable only at the national level and currently only in the Netherlands, as mentioned during the Dutch focus group. Time series are generally on an annual basis, thus

limiting applicable time scales. It is also considered that data aggregation reduces granularity and limits the possible analyses that can be applied.

Another interesting comparison that cannot be made based on existing data is exploring the differences in development between farms in flat, semi-mountainous, and mountainous areas. The regional specifics at the farm level are not generally available in rural areas. In almost all MSs, the data for statistical analysis at the rural area level is quite limited and scarce, except for the national censuses, which are carried out at longer-period frequencies. Observed social evolution limits opportunities for future development, remaining largely concealed.

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?

Standard statistical descriptive tools and data analysis methods are widespread, relatively easy to use, and can be quickly learned and mastered. They do not require any special degree or sophisticated expertise for practical implementation. The use of software products, such as Excel and SPSS, makes the implementation of statistical tools accessible and widely applied for analytical and decision-making purposes at various levels. Furthermore, although some statistical tools may not be well-known or widely used, there is sufficient technical documentation and tutorials available, making it relatively easy to become acquainted with their specific features (Peck, R., et al., 2008). A large part of the methods is even available in prepared software products (Excel, SPSS, etc.), which makes their application relatively accessible (OECD, 2008). This perspective on statistical methods was also confirmed during the Latvian focus group.

More skills and training, however, can be necessary for choosing and setting up the statistical method (e.g. often these tools draw upon economic theory and related assumptions, which need to be known and considered), and especially for interpreting the results (e.g. due to the economic assumptions and statistical facets). Yet, for simpler analyses the interpretation of results can be relatively straightforward (Killen, P., 2007), especially in cases where scientific and academic organisations are involved in 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 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?

Standard statistical tools for data analysis and descriptive statistics are available in software products such as Excel, SPSS, etc. (Arsham, H., 2020). Some software are open-access, while other are not. Descriptive statistical tools are often relatively quick and easy to find and apply, even without using specialised software. These tools are well-known, widespread, and widely used, as has been demonstrated in various cases and types of analyses and decision-making initiatives, offering numerous advantages that make them useful for researchers and practitioners. The application of those can allow a better understanding of characteristics, behaviours, attitudes, and beliefs on a particular phenomenon or population, allowing researchers to provide a detailed and accurate description and interpretation (Peck, R., et al., 2008). Descriptive statistics also allows the generation of hypotheses or ideas, by linking the potential causes or factors influencing a particular phenomenon or population can be scrutinised and verified, which can be useful in guiding future research. More advanced statistical techniques might not be available

in pre-build software packages and may require skilled professionals for the set-up. Some other statistical tools may be tailored to specific contexts and need to be re-adapted. For example, according to the evidence of the Dutch focus group, the IOI matrix is specifically designed for the Netherlands and is not easily applicable to other MSs.

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 sub-category of tools (i.e time necessary to organise and prepare the implementation)? And how much time does it take to implement the process of the tool (i.e. time necessary to implement the tool until the achievement of the intended outputs)?

According to our survey results, overall, the set-up of these tools requires less that one month, around 1 and 2 weeks. The time efficiency of standard statistical tools has been also acknowledged by the participants of the Dutch focus group. Beneficial is the availability of software products, which makes the task of data processing relatively quick and simple (Arsham, H., 2020). It should be noted that the time efficiency of using standard statistical tools for data processing and deriving results is quite high. As for data collection and preparation, it varies depending on different circumstances, but it is convenient that the capture of data in calculation tables can be done relatively easily and quickly (Peck, R., et al., 2008). More time can be required for preparation if ad-hoc data collection has to be carried out. However, this is rarely the case as several data sources, including official statistics, are available. Their ease of use, flexibility, and scalability also improve the time efficiency for end users, of all skill levels and sophistication.

Cost-effectiveness

Cost-effectiveness refers to the extent to which the tool is able to achieve its outcomes within a reasonable amount of 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 budget can be necessary for the maintenance of the tool or subscription for patent-based tools every year?

Standard statistical tools are relatively cheap and affordable compared to customised software products for more complex simulation models, making the cost-efficiency another strength of these tools (Abatan, S.M., Olayemi, M., 2014). The implementation costs account for expenses related to software or hardware acquisition, training, initial configuration or customisation, as well as maintenance costs for updates, hardware maintenance, and any necessary support services, along with subscription or licensing costs. These costs, however, are quite limited and affordable compared to other tools, being more cost-effective than other computation methods. Open-source tools are also widely available.

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 professionalities are necessary for the good implementation of the tool?

The number of professionals involved in the analysis can vary depending on the complexity of the analysis. However, it is not expected to require a huge number of human resources, at least compared more complex simulation models. Nevertheless, in several cases, as mentioned during the Dutch focus group, the expertise required can vary significantly. Different tools may require input from individuals with diverse backgrounds, such as technical experts, project managers, and data analysts, and cross-functional collaboration should be also considered. For a

comprehensive assessment, it is crucial to involve domain experts who possess expertise in the area under study (Koop, G., 2013). Our survey results indicates that usually less than five professionals are required, whereas these tools can require up to 50 to 100 hours dedicated to trainings.

3.3.3. Simulation models tools

Summary

***Accuracy**: moderate to high, as they facilitate traceability of policy choices, implement data verification measures, while being sensitive to changes in input parameters*

Simulation models' tools facilitate traceability of policy choices back to their origins and implement measures such as verification of statistical data to avoid mistakes and ensure consistency. Macro-level models also prioritise data reliability, constructing databases from reputable sources. Although tracing information back may be complex, these tools implement measures like data verification and regular updates to maintain data accuracy and validity throughout their usage. These tools are sensitive to changes in input parameters, particularly price-cost ratios. They can capture diverse facets of a system by modeling different farm types and structures.

***Reliability**: moderate to high, as they can replicate results, while different models may yield divergent results due to rigid assumptions. Their reliability varies with the accessibility for external review and the availability of documentation.*

Simulation models tools are sensitive to input changes and can replicate results under the same conditions. They capture diverse system facets and can be combined with macro-models for broader analysis. However, they can generate divergent results for the same scenarios, and can build upon rigid assumptions. These models generally undergo quality reviews and tests, and their model documentation is usually available. However, the accessibility for external review varies. For many models (e.g. SiTFarm, CAPRI, MAGNET) methodologies are known and accessible through published reports and can be accessible for external review, with training materials offered online. Some study results might be classified as sensitive and not available for external review.

***Applicability**: moderate, they are adaptable for specific objectives and used in strategic planning, but vary in their applicability due to differences in tool features and user expertise, yet they support multi-level, temporal policy analysis.*

Simulation models tools can be used to assess multiple policy objectives and instruments. They have been used in strategic planning and can be tailored for specific objectives. However, some tools (like SiTFarm) are static and do not allow analyses from the perspective of structural changes. Experience with Simulation models' tools varies by MS, but most have some expertise. Macro-models are frequently used by the EC for different policies. Currently, only a few MSs have the capacity to use these macro-models for country-focused analysis. These tools can support analysis and policymaking at different geographical levels and time scales.

***Accessibility**: limited to moderate, requiring specific expertise and training, they are managed by research institutes which support knowledge exchange, but their accessibility varies due to the need for local adaptation and advanced skills.*

Simulation models tools often require specific expertise, considerable training, and experience. They are typically managed by research institutes and require understanding of detailed data sources and mathematical models. For some of these tools, extensive trainings are provided by the research network. Despite their complexity, these tools support knowledge generation and exchange within their user communities. Availability of Simulation models' tools varies. Some tools (like Farmdyn) have a generic application around the EU but require local adaptation. National models may be usable in other contexts but require specific adaptations as well.

***Efficiency**: limited to moderate, while these tools are time-consuming and costly to set up due to their complexity, their efficiency is enhanced by the research network support, and once developed, they are relatively simple to use with costs mainly related to researcher time and licensing.*

Setting up and implementing simulation models tools can be time-consuming due to their complexity. The time required varies from weeks to months, or even years for new teams. However, these tools are part of a network, allowing users to address technical questions to experts in the field. Implementing these tools can be expensive, with costs mainly related to their development, adaptations, and maintenance. Once developed and adjusted, the use of these tools is relatively simple and the costs are not high, being mainly related to the working time of the researchers and the cost of the license. Such tools can be run by one researcher but usually involve a team of up to 5 researchers. While one person can run a model, a team of experts is needed to validate the wide-ranging results it produces.

Analysis by criteria

ACCURACY

Accuracy analysis considers how successful is the tool 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 back the information used for supporting the policy choices 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?

Overall, farm programming models that use PMP incorporate an artificial cost curve to precisely replicate observed farmer decisions in baseline scenarios. These models seek to optimise a goal function subject to constraints, but since they may not fully mirror real-world conditions, there can be discrepancies between model predictions and actual outcomes. The advantage of such models lies in their ability to trace results back to the model's functioning and understand the underlying mechanisms and behavioural adjustments (Agricultural Institute of Slovenia, 2023).

Model calculations serve as a foundational simulation tool, dynamically simulating agricultural production across different conditions. This system, through its built-in dependencies and real-time adjustments, becomes a primary source of analytical and economic data for individual production activities. The tool's structured organisation allows for tracing the impact of CAP measures on economic indicators at various levels, ensuring model verification against statistical and other available data.

The Slovenian farm model tool (SiTFarm)[50] exemplifies a complex simulation model-based tool for analysis at various agricultural levels. Its modular design, where the different modules take the form of MS Excel spreadsheets are developed to enable their independent operation or inclusion in a common integrated system. This linking of modules is made possible through several Visual Basic loops and ensures consistency. This feature is pivotal for conducting sectoral analyses from farm-level results, simplifying the aggregation process to reflect Slovenian agriculture comprehensively, although with necessary simplifications. The tool's integration with model calculations from the Agricultural Institute of Slovenia forms a critical simulation tool for economic indicators in agriculture (Žgajnar J., et al. 2022). The participants to our Slovenian focus group considered the accuracy of this tool as very good, although it needs continuous updates, and reported that data validation measures are in place. However, they acknowledged that the model has its limitations. Improvements in Slovenia's FADN, currently hindered by sample selection and data quality issues, could enhance assessments. During the Dutch focus group, for example, participants reported how the farm income models relying on FADN data offer insights based on a sample that accurately reflects the farm structure. However, it was also mentioned how adjustments to farmer behaviour in response to policy changes are taken into account in a limited way. Similarly, for the Eco-scheme farm simulation tool, which provides an accurate overview of received payments under a certain eco-scheme choice, it was

[50] https://www.kis.si/en/Model_Calculations_OEK/

mentioned that since it is a deterministic tool, it does not take into account uncertainty in case of possible oversigning effects.

In the case of models operating at the macro level, so covering multiple countries, such as CAPRI[51], the sources of data 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. Tracing back to original data sources can be complex and time-intensive, underscoring the need for technical expertise and careful data input validation to avoid errors and ensure data consistency (Bedrač M. et al. 2023; Žgajnar J., et al. 2022). 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. This long-standing commitment in providing quality statistical data has been acknowledged also from the participant of the Hungarian focus group.

The survey results indicate that farm programming models are strongly supported by well-founded evidence, as evidenced by the high ratings given by respondents: five respondents provide a score of 5/5, and two respondents gave a score of 4/5. Additionally, the majority of respondents reported that the process of data input is traceable, underscoring the reliability of the data used. However, there's uncertainty about measures to prevent errors and redundancy in data input, except in the Slovak Republic and Belgium (Wallonia), where the models used applies this kind of measures.

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 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 it seeks to represent?

Programming models of farms are valuable because they can simulate farmers' reactions to policy changes, side conditions, or shifts in key parameters, providing essential insights into their behavioural responses. The models' representativeness hinges on various factors, including the diversity and number of farms modelled, which reflects the heterogeneity of farmers' preferences, farm types, structures, and soil conditions. The degree of representativeness achieved can vary, often depending on the specific research question, as well as the available capacity and budget.

In Slovenia, the SiTFarm tool utilises simulation models, specifically linear programming (LP) for Gross Margin (GM) maximisation, sensitive to input parameter changes such as price-cost ratios. These changes can affect the economics and attractiveness of production activities. Challenges also arise with changing CAP measures and their conditions. The tool employs a partial optimisation process, where certain farm-defining variables (like livestock numbers or crop rotation ratios) are fixed, while others are variable and subject to optimisation. However, the baseline analysis aims not at classic optimisation for maximum GM but at reconstructing a practical production plan, often resulting in a compromise between various goals and conditions rather than economic maximisation. The optimisation primarily aims to balance material flows on the farm, producing a technologically feasible plan that imitates real situations and allows for analyses using technological, economic, and environmental indicators, although it may not be economically optimal. Currently, SiTFarm operates under deterministic assumptions without incorporating input variability, allowing for monitoring of income, variable costs, and gross margin at both aggregate and individual farm levels, representing diverse agricultural holding types (Žgajnar, J., et al., 2022)

Macro-models like CAPRI[52] can capture the diversity in regional impacts or across products and activities due to a high level of disaggregation. The sensitivity of the models, or how much the results change with variations in input data, is generally limited when only minor adjustments are made, due to the models' need to balance a vast array

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

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

of inputs, outputs, and intermediate consumption while meeting specific constraints and requirements. The CAPRI model, in particular, is designed to reflect the impacts of certain input data variations. Feedback from survey respondents indicates that overall the proposed tools have a high degree of sensitivity, showing significant responsiveness in the output to changes in input parameters or assumptions.

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?

In simulation models, replicability is considered an indicator of quality models and practices. Unlike other methods that focus on observed behaviours, past trends, or omit explicit optimisation, programming models are distinctive for incorporating farmer responses. This feature becomes particularly valuable in the context of new policy measures, where past information offers limited insight into future outcomes.

SiTFarm is a Simulation model tool that employs a static deterministic approach, ensuring reproducibility, where identical input conditions yield the same solution. Errors in results may arise from incorrect data entry, given the extensive system of input data required. To mitigate this, a detailed input module has been developed to prevent such errors. Nonetheless, due to SiTFarm's unique nature, its results cannot be replicated exactly with other available tools.

Replication of results of macro-models (like CAPRI or MAGNET) should not be the goal of other tools or through different approaches as each of the tools should serve the particular purpose it is built for. However, it is possible and, in many cases, useful the combination/linkage of tools to ensure broader coverage of questions to answer by using harmonised assumptions. As mentioned during the Hungarian focus group, the CAPRI model is run on different computers. Repeating the same scenario should result in only negligible discrepancies, since modifying scenarios would, therefore, modify outcomes as well. Additionally, setting up a scenario and writing the corresponding equations results in some unavoidable personal preference in the code. CAPRI[53] is covering global agricultural trade, but does not have a capacity to simulate macro-indicators, it can be combined with "real" macro level model (MAGNET for example) to use the same assumptions for macroeconomic development in order to provide analysis complementary to MAGNET results. Otherwise having structural differences in terms of the data and behavioural elements partial equilibrium (PE) and computable general equilibrium (CGE) models can generate divergent results or even contradictory findings for the same scenarios (Phillippididis, 2017). Although this is well recognised within the modelling community, in the policy arena it can often be hard to reconcile the findings of both models when presenting a consistent story line for a given policy reform. Several projects commissioned by European Commission forged a "soft" model linkage such that both models generate a mutually consistent storyline whilst one plays to the strengths of each model to serve as a source of input to the other.

In terms of replicability, it may be mentioned that a specific project, the Stable Release (STAR) project, had been launched to testing and improve the coherence and stability of the entire system. In this project the model was run with different settings or with different starting values while comparing the results in systematic way, to make sure that results may be replicated if the system is installed in different agencies.

[53] https://www.capri-model.org/dokuwiki_help/doku.php?id=scenario_simulation#sensitivity_analysis

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 simulation models, it is considered good modelling practice for tools used in policy assessment to conduct review and testing to ensure their quality. Furthermore, the robustness of model results is typically verified by a supervisory committee or guidance group within the context of a project, as well as by an internal quality review conducted by the research institute performing the analysis. Established programming tools typically come with documentation accessible to the public. However, survey respondents report a lack of availability or uncertainty about the accessibility of documentation detailing the tools' underlying methodologies.

For instance, the SiTFarm model is not an open-access tool that is freely available, making external review in such a manner impossible. However, the underlying methodologies are known and can be accessed through various published reports. The input data (economic, technological) of the tool largely relies on model calculations (MC)[54], which, to a certain extent, are accessible and published on a monthly basis by the Agricultural Institute of Slovenia. These include reference budget calculations, which may differ from those used in SiTFarm due to adjustments (e.g., the amount of milk yield per cow, crop yields, etc.), but the processing and production technology is clearly outlined, including the cost ratio. Other farm simulation tools are available. During the focus group conducted in the Netherlands, it was mentioned that the Eco-scheme farm simulation tool is available online for users, including farmers, and directly provides information related to their choices.

The information regarding the underlying methodologies and data of macro-models like CAPRI and MAGNET is readily available and accessible, and the core version of these tools is open for external review. Similarly, as mentioned also during the Hungarian focus group, all the linkages with well-established and harmonised databases is also well documented, increasing then the transparency of the tools. While this requires a certain level of expertise in the modelling field, training materials are provided on the website, and training sessions are available for those who register. Although the input information 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.

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?

General programming tools for farms can assess multiple objectives and policy instruments. They can also provide insights into potential interaction effects at the farm level resulting from the simultaneous implementation of different policy instruments. Moreover, they can offer information on various types of impacts, such as economic impacts (income, costs), agronomic (herd management, crop rotation, fertiliser/manure application, etc.), and environmental

[54] https://www.kis.si/en/Model_Calculations_OEK/

indicators (e.g., CO2 and ammonia emissions). Farmdyn and CAPMAT serve as prime examples of such tools, known for their broad coverage (Alif, Ž., Šumrada, T., Žgajnar, J., 2023; Brečko, J., Žgajnar, J., 2022; Pečnik, Ž., Žgajnar, J., 2022).

The primary purpose of the SiTFarm tool is to facilitate the assessment of impacts from the perspective of income sustainability. Therefore, the model calculates various economic indicators and allows for the inclusion of various CAP measures under different levels and conditions (socio-economic context of analysis). This tool was also utilised in preparing the CSP for Slovenia (Žgajnar, J., Juvančič, L., Kavčič, S., et al., 2021). During the preparation phase, a larger number of scenarios were implemented and tested (ex-ante analysis), enabling policy support in decision-making processes (intervention setting, budget allocation, and stakeholder consultations) (Brečko, J., Žgajnar, J., 2023; Žgajnar, J., Zadnik, Stirn L., 2023), but also It can be used for monitoring, legitimation, budgetary planning and agricultural outlooks and choice of intervention or assessment of potential new interventions, as mentioned during the Slovenian focus group.

Besides economic indicators, the model also provides some environmental indicators, offering insights from an environmental perspective. However, it should be noted that the model does not support analyses from the perspective of structural changes or impacts on structural changes, as it is static in this regard (El Benni, N., et al., 2023).

SiTFarm has also been applied to analyse biodiversity perspectives and could be further developed in this regard. Specifically, it has been used to analyse the impact of price-cost ratios on business economics and on economics at the level of certain types of farms, to assess various environmental indicators at the farm and sector levels, to analyse the impact of introducing nature conservation measures, and to determine farm economic viability in the case of a Slovenian Sub-Mediterranean region facing land abandonment (Alif, Ž., Šumrada, T., Žgajnar, J., 2023). Another application involved analysing risk and risk management strategies for selected farm types (Brečko, J., Žgajnar, J., 2022). Based on these cases, the model can be adapted or upgraded to include these aspects as well.

The CAPRI model is well-suited for evaluating the impact of the CAP, trade, and environmental policies on agricultural production, income, markets, trade, and the environment, on both a global and regional (NUTS2) scale. As mentioned also during our Hungarian focus group, the tool is designed to focus on the first Pillar of the CAP, encompassing both coupled and decoupled subsidies. It also addresses significant subsidies from the second Pillar, including support for Less Favoured Areas, agri-environmental measures, and Natura 2000 support. Thanks to its detailed spatial resolution at the NUTS2 level, the tool is adept at simulating the effects of decoupled payments. However, within the CAPRI model framework, any payment scheme in a MS is essentially an approximation. The model, which is based on activities, faces challenges in accurately implementing subsidy schemes that do not require production, leading to less precise results in scenario analyses. Nevertheless, 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[55] and it can be adjusted to assess and contribute to the assessment of specific objectives in foresight, horizon scanning, and impact assessments.

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?

The experience in programming farm models may vary by MS, but in most of them, considerable expertise is likely to exist. This is confirmed by the survey findings, which indicate that some of these tools were utilised in the previous CAP programming period, notably in Poland, the Slovak Republic, and for SiTFarm in Slovenia. The experience specificity for each MS is particularly true as programming (e.g., linear programming) has been a widely used tool

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

at the agronomic-economic nexus to understand farm behaviour and policies and this kind of knowledge should be available especially in academia. However, there might be constraints in replicating or transferring this experience from one country to another, due to its specificity. Similarly, also the experience of simulation model has been confirmed during the Latvian focus group, where participant highlighted how these tools were used in the design and planning of some interventions and assessing the impact of changes in support rates on result indicators.

This could be the case of SiTFarm model, which is a tool adapted to conditions in Slovenia, making it unsuitable for direct use in other countries[56]. The transfer process is currently underway in the dairy sector, which should enable its use in Croatia as well. From the conducted survey was also clarified that this tool has been used in different policy areas other than the CAP. On the other hand, this could be the case also macro-models, like CAPRI, which are frequently used by various Commission services, such as DG AGRI, DG ENV, DG CLIMA, and JRC, for reporting on agricultural, environmental, and climate policies at the regional dimension in the EU[57]. Since July 2017, this model has supported the ex-ante impact assessments of the European Commission. So far, only a few MS agencies (Germany, Spain, and Sweden) have the capacity to use the model for country-focused analysis. Tools4CAP presents an opportunity to test whether the required expertise is transferable within a reasonable amount of time.

Macro-models have been extensively used, especially by the EU, to assess EU policy, policy reforms, and trade, as well as trade-policy issues. Recent applications include the assessment of the impact of COVID-19 and the Ukraine-Russia war (Jongeneel, R. et al., 2022). But also, several ex-ante assessments have been done, which have been useful as inspiration for future policy reforms (e.g., climate scenarios in ECAMPA projects) (Perez Dominguez, I. et al., 2016).

Numerous publications cite the use of the CAPRI model in various aspects of agricultural policy planning in the EU. It is accessible without any physical barriers, offered free of charge, and designed with policymakers in mind. However, as mentioned during the Hungarian focus group, its wider adoption by the agroeconomic community is limited due to the requirement for an in-depth understanding of its code structure to implement custom scenarios. CAPRI provides results at a regional level for economic and environmental variables, allowing for economic and environmental analysis of different policy scenarios regarding reforms of the CAP. It can perform regional-level analyses of specific Common Market Organisations (e.g., sugar, dairies), trade of agricultural goods with the rest of the world (e.g., WTO proposals), environmental policies (e.g., greening, climate action, and water), and different subsidy schemes in Europe (e.g., partial decoupling of agricultural subsidies). Operational for two decades, it supports decision-making related to the CAP of the EU and its implementation at the MS and regional levels. The effects of agricultural, environmental, and trade policies on agricultural production, consumption, prices, and trade, as well as on environmental indicators, are analysed in a comparative-static framework, where the simulated results are compared to a baseline scenario.

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)?

Programming models at the farm level can focus on individual farms as well as on farms that are representative of a certain area or region (Žgajnar, J., Kavčič, S., Tomšič, M., et al. 2022). Therefore, different granularity types can be used, which may have implications for the analytical focus. For example, a regionally representative farm might be expected to be less specialised and include a wider regional crop mix than the actual individual farms in that region. The choice of granularity and focus will depend on the research question, and which is best suited for the purpose.

Overall the applicability of simulation models is mainly at national level, as mentioned during the Latvian focus group. In the case of the SiTFarm model, for example, this provides mainly results at national level, with some of the

[56] https://www.mpsr.sk/strategie-analyzy-a-prierezove-cinnosti/rezortna-statistika-od-2020/47-242-1523

[57] https://www.mpsr.sk/strategie-analyzy-a-prierezove-cinnosti/rezortna-statistika-od-2020/47-242-1523

required data is collected on a regional and/or local basis. Each sector has a different timescale for data collection (Žgajnar, J., Kavčič, S., Tomšič, M., et al. 2022). The possibility of being applied at smaller or sub-national scales and the principle to be applied at higher spatial/governance scales was also mentioned during the Slovenian focus group. In the future, it will be possible to perform analyses at lower levels, but this step will require adjustments to the model and additional data. From a temporal resolution perspective, it is possible to conduct analyses over different periods (Žgajnar, J., Kavčič, S., Tomšič, M., Hiti Dvoršak, A., et al., 2021). Currently, the price list is prepared and adjusted for the periods from 2017 to 2022. In support of strategic planning, the analyses were based on three-year periods (2017-2019, 2018-2020, 2019-2021, and 2020-2022). Additionally, analyses can also be performed for shorter periods – half-yearly or monthly within the period from 2017 to 2022. In the case of CAPRI, it provides results for trade at the country or region level, while results for agricultural production can be downscaled to the NUTS2 level (Žgajnar, J., et al., 2019). This was also confirmed during the Hungarian focus group, where it was mentioned that units allow for aggregation, enabling CAPRI to calculate national and EU-27 results.

The periodicity of macro-models is usually year-by-year, sometimes even with intervals (e.g., 5-year periods). For time scale consideration, the dynamics of a model are important. Statistical/econometric models often have rich dynamics (including past and/or lagged variables) relative to macro-models, which have a comparative static nature. The latter compare “old” and “new” economic equilibrium situations, whereas the former can also provide information about the adjustment process over time, needed to arrive at a new equilibrium state. Several macro-models (e.g., MAGNET, CAPRI) now contain some form of modularity, where specific modalities of the model can be used optionally. This includes regional (de)aggregation.

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?

Maintaining and running farm programming models is typically the responsibility of research institutes, as they possess the necessary manpower and investment capabilities to manage such tools. Often, specific expertise is required to operate these models and to translate policy signals into model parameter adjustments. From the survey this has been confirmed, with the majority of respondents indicating that the use of the proposed tools requires previous experience.

This is also the case with the SiTFarm tool, which is not open access. Due to its complexity, the model requires considerable training time and experience with mathematical models and this was also confirmed during the focus group conducted in Slovenia. Even though the model allows analysis at the sectoral level, it is supported by many very detailed data sources (at the production activities level), and knowledge of these, as well as the place of influence, is important for the correct use of the model. A sufficiently powerful solver for Simulation model is also necessary.

Similarly, macro-models generally have a significant degree of complexity, and researchers need to be well-trained and experienced to use and apply these tools. A prerequisite, especially for macro models, is that timely investment in the research/model infrastructure is desirable to enable the use of the tool at a specific moment with relative ease. Often, this type of models is hosted in one or more research institutes and are used, maintained, and extended by a community of researchers. Examples includes the GTAP community based in Purdue, the CAPRI community with “focal points” in Bonn and JRC Sevilla, the AGMEMOD community with core partners in The Hague (WecR) and Braunschweig (Thuenen Institute), together with supporting partners from other MSs such as Ireland, Poland, Italy, and France. These research communities have the advantage of being able to disseminate learning experiences and share maintenance, as well as investments. Moreover, local researchers can benefit from the international

research community for advice, help, and support in case of the application of the macro-tool at national level. In the survey has also been reported from the respondents that for this type of tools detailed documentation concerning the tool's underlying methodology, models, algorithms, and analytical techniques is not available and accessible to all users.

Using CAPRI, for example, requires extensive knowledge of how the system works. It demands a certain level of expertise, prior training, and experience, as well as time to become proficient in its use. This was also highlighted during the Hungarian focus group. The CAPRI network organises annual CAPRI training[58] with the GAMS pre-course (3 days training for the basics of the simulation models language GAMS used in CAPRI), a CAPRI training session for beginners (5 days), and a CAPRI developer meeting considered as advanced training on different topics and modules built and analysed by CAPRI. The CAPRI training aims to bring together modellers and developers with a strong interest and background in agriculture-related research to improve their knowledge about agricultural economics and quantitative modelling of agricultural, environmental, and trade policies with the CAPRI system. Single bare-bone instructions in documentation may not be sufficient to get the system up and technically running. CAPRI Stable Release (STAR) versions used for training purposes are published with a full set of data, i.e., including all the intermediate data required to build the complete database and produce a calibrated baseline. It can be handled by economic modellers with programming skills considered as an important asset. Discussions on common issues on agricultural policy, policy modelling, model development, and maintenance with CAPRI developers and users support knowledge generation and exchange of experiences in the community (European Commission, 2021).

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 survey reveals uncertainty or a lack of availability of the proposed tools for end-users across various MSs. However, it identifies programming tools like Farmdyn with potential for EU-wide application, although necessitating local adaptations to align with specific policy environments and soil conditions.

In Slovakia, for example, results are publicly available, but methodological and technical guidelines are not. In Slovenia, the SiTFarm model is not openly available to everyone. The main reason is its complexity, as it integrates several modules and a tool (MC) from the Agricultural Institute, which is not publicly available. Making it openly available would also entail high costs and likely reduce flexibility in analyses. On the other hand, macro-models like CAPRI are available as open source for end users. Nevertheless, as reported by the participants of the Hungarian focus group, the basic equations are embedded in the core part of the CAPRI model and are accessible only to code maintainers with long-term expertise. This means that CAPRI has limited accessibility in this respect.

Standards are available that can enhance the usefulness and dependability of agricultural models in various scenarios. This includes the “Quality Indicators for External Users” provided by the FAO[59], which underscores the critical role of high-quality and reliable statistical data in the effectiveness of agricultural models for policymaking. It provides detailed guidelines on essential quality indicators such as relevance, accuracy, timeliness, and data accessibility.

Access to macro-models and their research communities is generally quite open. Nevertheless, in the conducted survey has been indicated that the tool is not available to tool available for end-users in all MSs. A user who is interested in the application of the tool often has to sign a Memorandum of Understanding (MOU) to become a member and gain access to the tool, including model code or data, and become a member of the modelling community. There is some reciprocity involved: the receiving partner can benefit from being able to use the tool and participate in provided training. In return, it is expected that the new country team contributes to the model consortium. This contribution may be in financial terms, for example through projects, or in kind, by providing data

[58] https://www.capri-model.org/ts/dokuwiki/doku.php?id=start

[59] https://www.fao.org/3/cb9309en/cb9309en.pdf

on the national situation and contributing to model improvement. Macro-models have fixed costs and require continuous extensions, updates, and improvements, which are costly, especially because such models usually cover a wide range of countries.

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 sub-category of tools (i.e time necessary to organise and prepare the implementation)? And how much time does it take to implement the process of the tool (i.e. time necessary to implement the tool until the achievement of the intended outputs)?

Programming models vary in terms of detail and complexity, which translates into labour and expertise requirements for operating such tools. The time required to set up the proposed tools may range from weeks to months, depending on the model's detail and the research question, and this has been confirmed also from the survey evidence.

For instance, as mentioned during the Slovenian focus group, the SiTFarm tool requires constant updating with new data on production conditions, technologies, and farm structure to remain relevant. It was pointed out that many models become obsolete if they are not constantly used and updated, which requires significant resources. The tool allows for customisation and modification at various analysis levels. Some changes and analyses are relatively simple (e.g., analysis of price-cost ratios, changes in payment levels, minor adjustments in the conditions of certain policy measures), while others are more demanding (e.g., changes in production technologies, introduction of new measures, addition of new farm types to the aggregate, introduction of new production activities, etc.). The time required for analysis also varies. However, from the perspective of time requirements, it could be classified as a more time-demanding tool.

According to the findings from the Hungarian focus group, overall the CAPRI model is more time efficient compared to other tools. However, the set up of large or macro-models scenarios with CAPRI can be time-consuming, requiring a significant amount of time. Establishing a new modelling team in an agency that has not previously used CAPRI can take up to a couple of years to set up and implement the process with significant assistance from core developers. Researchers will also need sufficient capacity to focus on the system and become fully independent. However, as the system is used by an entire network, it will be possible to address technical questions to experts in a specific area.

The setup for simulations of farm income is relatively fast, taking only a few days to weeks, as noted in both the Dutch and French focus groups. Regarding the Eco-scheme farm simulation tool, discussions during the Dutch focus group highlighted that while the development and maintenance of the web-based tool are time-consuming, the tool itself is user-friendly, requiring only a few hours to a couple of days for users to become proficient.

COST-EFFECTIVENESS

Cost-effectiveness refers to the extent to which the tool is able to achieve its outcomes within a reasonable amount of 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 budget can be necessary for the maintenance of the tool or subscription for patent-based tools every year?

Programming models may require a significant budget, especially related to their setup, but also for application, maintenance, and reporting of their results. Assuming the model is in good foundational shape, for specific project applications, the budget may start at 30,000 Euro and increase depending on the type of research question and the approach taken, potentially reaching 50,000 Euro, as indicated by some respondents during the survey.

It is typical for models of this type, like SiTFarm in Slovenia, that the key costs are related to their development, adaptations, and maintenance. Once the model is developed and adjusted to the appropriate conditions, its use becomes relatively simple, and the costs of such analyses are not high. Costs are mainly related to the working time of the researchers and the cost of the license for the solver. Similarly, macro-models like CAPRI are expensive primarily because of personnel costs, as they require staff who are highly skilled and specialised in this tool. Additionally, as also mentioned during the Hungarian focus group, a GAMS program license is necessary for running the model, representing the only significant investment cost needed.

Macro-models are essential for analysing a country or region within an international context, particularly for examining market interactions and trade relationships in scenarios such as trade reforms or geopolitical shocks. Despite their resource-intensive nature, their application is considered cost-effective for these specific research questions, as there are no alternatives that can effectively replace macro-models in such contexts (European Commission, 2023). Sometimes a fee must be paid for using them, but this is not always the case, although there normally will be a reciprocity condition for participation. According also to the survey results, it is not possible to make a precise estimation of the budget needed and costs for use and maintenance, as most of the macro-models are often designed for specific research projects, which enable model development and application. During the focus group conducted in Italy, participants mentioned that the simulation tool for elaborating scenarios concerning BISS and CRISS requires 2,000 Euro and the involvement of more than 5 people, due to the very specific skills required.

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?

The agricultural data landscape is rather complex and limited. Navigating this arena requires experience and knowledge of a broad spectrum of stakeholders and actors involved in their collection. It is important to have knowledge of who collects what data and for what purposes.

Programming models can be run by a single researcher but usually involve a team of 3 to 5 researchers, allowing for a division of tasks, and this has been confirmed also by the conducted survey. The SiTFarm model, for instance, can be very efficient. In general, one person is enough to operate, start up and use the model. However, since the model covers a very wide area of application, it is necessary to have a team of experts who can validate the results, with at least two researchers, depending on their expertise. Also, for the CAPRI model, the number of people required is not as crucial as the capacity, knowledge, and skills to work with this tool. During the Hungarian focus group, it was highlighted that the need for regular updates of the CAPRI system is constrained by a workforce bottleneck, resulting in less frequent major revisions than desired. Furthermore, access to the CAPRI model is limited to code maintainers with extensive expertise, underscoring the critical role of human expertise in coding complex scenarios within CAPRI.

The focus group conducted in Slovenia also suggested that the SiTFarm tool requires human resources to be aligned with the need for constant updates with new data on production conditions, technologies, and farm structure to remain relevant. In Italy, according to the focus group, the simulation tool for elaborating scenarios concerning BISS and CRISS required 2,000 Euro and the involvement of more than 5 people.

As previously indicated, macro-models are usually supported by an associated international research community. These communities are responsible for model development and maintenance, organising conferences, collaborating on projects, and as well for providing the necessary training. The effective utilisation of macro-models necessitates the integration of researchers into relevant communities, as these applications typically demand teamwork. Being part of such research communities offers the advantage of accessing and leveraging project-relevant experiences and expertise. Overall, as mentioned during the focus group conducted in Latvia, simulation tools require experts who have the knowledge and skills necessary to use such models. During the Dutch focus group, it was also mentioned that the Farm Income Model requires individuals with a good understanding of FADN, data infrastructure, and software/programming (SPSS).

3.3.4. Key findings

In this section, we summarise and cast light on the key aspects emerging from the evaluation of policy analysis tools for evidence-based decisions, and we conclude by identifying the strengths and weaknesses of this category of tools.

Policy analysis tools are used to evaluate the potential impacts of policy decisions, providing insights into their various effects. These tools facilitate decision-making by assessing multiple policy options, predicting outcomes, and informing policymakers about potential consequences and trade-offs.

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

Accuracy is assessed as moderate (for statistical analysis) or moderate-to-high (for simulation models tools). Statistical analysis tools use reliable data sources and descriptive statistics for accuracy but can face challenges due to statistical consolidation. Simulation models tools, on the other hand, ensure traceability of policy choices and consistency through measures like data verification and capture diverse system facets.

Reliability is judged as moderate-to-high for both sub-categories of tools. Statistical analysis tools offer moderate to high reliability due to their standardised nature and replicability, fostering consistent results when applied to the same data population. However, different methods may yield varying outcomes. Simulation models tools are also reliable, but sensitive to input changes and can replicate results under the same conditions. They capture diverse system facets but can generate divergent results for the same scenarios. Both tools undergo quality reviews and tests, but their accessibility for external review varies.

Applicability is assessed as moderate (simulation models) or moderate-to-high (statistical analysis). Statistical analysis tools are widely used for data collection, processing, and analysis, providing quantitative results and verifying decisions, and can be adapted to different purposes. Simulation models tools can assess multiple policy objectives and instruments and have been used in strategic planning in some MSs. Yet, they may require adaptations when applied in new contexts, which suggests moderate applicability given their current development.

Accessibility is judged as limited-to-moderate (simulation models) or moderate-to-high (statistical analysis). Statistical analysis tools are widely available and easier to use, but macro-models require more advanced skills and timely investment in research infrastructure. Simulation models tools often require specific expertise, considerable training, and experience. They are typically managed by research institutes and their availability varies. Some tools have a generic application but require a local application, while others are not openly available due to their complexity and associated costs.

Efficiency is judged as limited-to-moderate (simulation models) or moderate-to-high (statistical analysis). Statistical analysis relies on quick setup and time-efficient data processing. They are cost-effective due to standardisation and software availability. Simulation models tools also have a more limited efficiency. They can be time-consuming to set up and implement due to their complexity. While their implementation can be expensive, their use is relatively simple once developed and adjusted.

Policy analysis tools for evidence-based decisions are crucial to assess the effects of policy decisions and can play a key role in policy tasks like needs assessment, interventions and target setting, or ex-ante analyses. In some MSs, specific policy analysis tools were developed to assess particular interventions, whereby their general accuracy and reliability can help inform policy decisions. Yet, many of these tools suffer from limited accessibility in some cases, high costs, time and skills, and often require necessary adaptations when applied to new contexts (e.g. when implemented in MSs for the first time). While the crucial role played by policy analysis in CSP design has become clear, these limitations underscore the need for timely investments, peer-to-peer learning among MSs, and reliance on the EC and/or existing networks of experts.

3.4. Monitoring and data collection tools

3.4.1. Inventory of monitoring and data collection tools

Table 9 below shows examples of tools belonging to the different sub-categories, according to our inventory of tools. The full list of tools is available in the Tools4CAP online inventory.

Table 9. Sub-categories of tools and examples of tools from the Tools4CAP inventory.

Sub-category of tools	Example of tools identified
Compliance monitoring	• Area Monitoring System (AMS) (Czechia, Germany, Netherlands, Malta, Czechia) • Digital Farm Book (Spain) • Classyfarm (Italy) • iSIP – Online Parcel Identification System (Portugal) • 3STR – Monitoring (France) • Bedja Cam (Malta) • Land Parcel Identification System (LPIS) (Malta, Czechia) • Checks by Monitoring (CbM)
Performance monitoring	• New Monitoring System AGEA (DATA Platform) (Italy) • Irrigation Calendar (Calendário de Rega) (Portugal) • Digital Farm Book (Spain)
Data and knowledge stocktaking	• Digital Farm Book (Spain, Portugal) • Farm Accountancy Data Network (FADN) • Integrated Administration and Control System (IACS) • Area Monitoring System (AMS) • Land Parcel Identification System (LPIS) (Malta, Czechia)

Monitoring tools have been employed to support different tasks. Table 10 below provides an overview of the number of MSs where a specific sub-category of tools was employed for each task of the CSP design and monitoring. More details are available in the online inventory of tools[60].

Table 10. Use of monitoring tools across the CSP design and monitoring (no. of MSs in which the tools were employed for the specific task).

	Compliance monitoring tools	Performance monitoring tools	Data and knowledge stocktaking tools
CSP design tasks			
Socio-economic context analysis			
SWOT analysis		1	
Needs assessment			
Interventions setting		1	
Target setting			
Financial allocations			
Ex-ante analysis and SEA			

[60] https://www.tools4cap.eu/tools/

Stakeholders' consultations			
CSP monitoring tasks			
Performance review		7	1
Beneficiaries' compliance	9		1
Evaluation	1		1

The following sections presents the results of the evaluation specific for each sub-category of tools.

3.4.2. Compliance monitoring tools

Summary

Accuracy: *high. Generally, compliance monitoring tools ensure data accuracy and validity through cross-checking and systematic quality checks.*

The data gathered by compliance monitoring tools are combined to verify that beneficiaries meet the requirements of the interventions from which they benefit, effectively creating a cross-check system to ensure data accuracy and validity. The tools are designed to collect only the essential information needed for this verification process, focusing specifically on interventions within the CAP and not extending data collection to other sectors or non-CAP agricultural holdings. Many of these tools use geotechnologies, like remote sensing for area-based interventions, capable of covering the full geographical extent necessary. To guarantee data quality, systematic quality checks are applied to these tools. The data, although not kept in their original form, are accurate and based on actual observations at both the single entity and farm levels.

Reliability*: high. Reliability and data transparency are ensured through data triangulation, which includes direct beneficiary information, off-farm data, and the national AMS.*

The reliability of these tools is ensured through the triangulation of various sources with some tools collecting information directly from beneficiaries and others using off-farm data acquisition technologies like remote sensing. These two types of data complement each other within the IACS to assess beneficiary compliance. The survey results indicate that the process is generally considered replicable, especially in the case of the AMS, which follows established methodologies and undergoes a Quality Assessment process for validation by the JRC. Compliance Monitoring systems, data, methodologies, and results are available and transparent.

Applicability: *moderate. These tools could contribute to most of the SOs and provide high-quality data collection for annual compliance verification to the CAP.*

Data provided from compliance monitoring tools can support multiple CAP objectives, ehnancing geographic area-based interventions and be effective in measuring compliance with environmentally related measures. These tools have also been used in previous programming periods and support analysis at the farm level, including IACS, which has been used exclusively in the CAP context. A new development for this programming period is the inclusion of the AMS, which is mandatory for all MSs. Since these tools are tailored for specific monitoring purposes, such as compliance, the level of detail they provide can vary. The information collected by these tools can be scaled up to different spatial levels. This can range from parcel-level information to geotargeted data, but overall, the granularity results as high quality. In terms of temporal resolution, as these systems are designed to verify beneficiaries' compliance during the current season, they typically deliver data once a year.

Accessibility: *limited to moderate. The tools require different expertise and have accessibility issues at the EU level, with some MS developing their own applications.*

Survey results suggest that there may be issues regarding broader accessibility at the EU level, with some MS developing their own applications tailored to their specific needs. The most common tool, the AMS falls under the category of IACS and geospatial monitoring systems but is currently not available to all stakeholders in the EU. Additionally, while remote sensing has been utilised for CAP control tasks (Check with Remote Sensing, CwRS), the AMS aims to exploit the capabilities of this data source fully. Effectively using this information requires expertise in data acquisition, processing, and integration into the IACS. Therefore, individuals with knowledge in various fields, including remote sensing, ITC development, and policy, are required. Training on interpreting the outputs of these systems is also essential.

Efficiency: *limited to moderate. The efficiency of these tools is limited due to the long time to set-up, costs are very high, several professional profiles needed from different fields of expertise.*

The time required to implement compliance monitoring tools, as well as the associated costs and expertise, varies significantly among different MS. Regarding the AMS, the estimated time frame for setup is based on Regulation (EU) 1306/2013, which establishes that MS shall have an operational AMS by 1 January 2024. According to the survey results, setting up the AMS would have taken one year, with more than six months needed for full operationalisation. In the case of those MS that opted for a gradual implementation of the tool, it would have taken three years, with the tool becoming fully operational by the time the regulation entered into force. The inventory does not provide insights into the budget invested in implementing the AMS, and it is unclear whether these organisations developed the tool in-house or relied on private companies. According to the survey results, in some cases the AMS requires an investment of more than 50,000 Euro and specific professional skills, involving between 50 to 250 people.

Analysis by criteria

ACCURACY

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

DATA ACCURACY AND VALIDITY

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

Evaluation Question: Are measures in place to avoid mistakes and redundancy and to ensure consistency in monitoring and/or the data collection activities (e.g. to verify the quality and reliability of the collected information)? Are collected data triangulated, corroborated, or substantiated by other data sources to check validity/accuracy? Are the collected data traceable back to its original source? To what extent is the data collected based on actual/direct observations rather than on estimations?

The main purpose of the tools under this category is to verify that beneficiaries meet the requirements of the interventions they are benefiting from. The information delivered by the tools under this category is combined to make a judgement on the beneficiaries' compliance. This combination could be seen as a sort of cross-check system to ensure the accuracy and validity of the data. Therefore, the information provided by these tools is accurate enough for its purpose. For instance, Earth Observation and geo-tagged photos are integral parts of the Area Monitoring System (AMS), that has been mentioned only by Spain within the survey. The AMS enables the assessment of the requisites of some area-based interventions[61]. Thus, the AMS, along with the Land Parcel Information System (LPIS), are two of the building blocks of the Integrated Administration and Control System (IACS). The information stored in the IACS may be combined with the information coming from different administrative databases and registries within the so-called Traffic Light Systems. The Traffic Light Systems deliver a judgment on the compliance (green light), non-compliance (red light), or non-conclusive (yellow light) state of each potential beneficiary applying for subsidies.

The conducted survey highlight that in most cases data is based on actual observations and that validation measures are in place, specifically data validation and verification processes, as well as standardised data collection protocols, with overall only 7 of the 14 respondents reporting that data are kept in its original form. For the validity of the data collected by these tools, the regulation lays down that "MSs should carry out an annual quality assessment of the identification system for agricultural parcels, of the geo-spatial application system and of the area monitoring system. MSs should also address any deficiencies and, if so, requested by the Commission, set up an action plan"[62]. Therefore, it is compulsory for MSs to implement the necessary quality assessment controls that ensure the accuracy of the data. In this regard, the JRC issued technical guidelines setting up the methodologies to assess the quality of the LPIS[63] and the Checks by Monitoring[64] (system upon which AMS is built).

[61] https://joint-research-centre.ec.europa.eu/system/files/2019-04/17-ams_dgagri.pdf

[62] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021R2116&qid=1702232434387

[63] https://cmapspublic3.ihmc.us/rid=1MW1WKXFS-1XYC88P-15GS/LPIS_QualityInspections_EURequirements.pdf

[64] https://publications.jrc.ec.europa.eu/repository/handle/JRC128276

TOOL REPRESENTATIVENESS

Representativeness refers to 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: Are the tools able to capture the diverse facets, variations, or elements within the system or dataset it seeks to represent? What is the potential geographical/sectorial/stakeholder coverage of these tools? Are these tools able to gather information from different sectors, stakeholders, and countries/regions?

The tools identified under this section are designed to collect information from all beneficiaries applying for CAP support with the purpose of checking the applicant's compliance with requirements of the intervention. Therefore, the tools only collect the necessary information to carry out this verification. Considering the focus of these tools, they are not expected to either collect data from other sectors or stakeholders nor from agriculture holdings outside the CAP. Some of the tools rely on geotechnologies such as remote sensing to collect the necessary information for area-based interventions. For instance, this technology, is able to collect information on the full geographical extension. In this regard, the geographical coverage of these data is ensured at European level. Moreover, the Regulation 2116/2021[65] encourages its use: "MSs should continue to use data or information products provided by the Copernicus programme, in addition to information technologies such as Galileo and EGNOS, in order to ensure that comprehensive and comparable data is available throughout the Union for the purposes of monitoring agri-environment-climate policy, […], and for the purposes of boosting the use of full, free and open data and information captured by Copernicus Sentinels satellites and services".

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 or other tools aimed at answering the same question, either by replicating the same approach or using new data or methods.

Evaluation Question: Can the data collected by these tools be replicated successfully by others or through different approaches? Do the data remain consistent with other data collection activities or tools? Does collected data remain consistent over multiple data collection rounds (i.e. the collected information is always the same if data collection is replicated)?

The tools in the inventory collect the required information through either direct interaction with the beneficiary (Digital Farm Book, Irrigation Calendar, for instance) or "off-farm" data acquisition technologies, such as remote sensing images, which help to reduce beneficiaries' interaction. These two types of information are complementary and combined within IACS to provide a consistent view on the beneficiary's compliance with the requirement of the intervention.

In terms of replicability, and as for the tools requiring beneficiary interaction, the information they collect may be subject to change either due to errors or modifications allowed in the regulation, but once the information is communicated to the system (IACS), the process of judging compliance should be replicable. As for the "off-farm" data acquisition technologies, they collect the information from external repositories maintained by trustworthy entities or organisations. For instance, remote sensing images from Sentinel-2 constellation are accessible from public repositories managed by ESA (European Space Agency), which keeps records of the raw information.

According to our survey respondents, in general, IACS and geospatial monitoring systems follows established methodologies and the results are subject to a specific Quality Assessment process designed by JRC for validation.

[65] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021R2116&qid=1702232434387

Moreover, the system allows for feedback and is kept up to date with the latest information available (the update of Sentinel imagery feeds the tool, so new images are obtained every 5 days for S2).

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?

The EC introduced the concept of Checks by Monitoring (CbM) in 2018[66]. The CbM is a verification method of beneficiaries' compliance with the legal bases of certain interventions and can be considered a first step towards the final implementation of AMS. The JRC issued, in 2018, the report "Technical guidance on the decision to go for substitution of OTSC by monitoring"[67] explaining the technical foundations of the new monitoring system. In 2020, the European Court of Auditors issued the report[68] "Using new imaging technologies to monitor the Common Agricultural Policy: steady progress overall, but slower for climate and environment monitoring" recommending a catalogue of best practices for CbM adoption. The JRC, based on the ECA's recommendations and on the experience of CbM early adopters, published in 2021 a technical report[69]defining, describing, illustrating and providing best practice considerations for the concepts on which the CbM should rely. Finally, in December 2021, the regulation[70] establishes that "MSs shall set up and operate an area monitoring system, which shall be operational from 1 January 2023". Although each MS has developed its own AMS adapted to its local conditions, the underlying methodologies and data are available and accessible since they are based on the above-mentioned technical documents.

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? Have these tools (i.e. their collected data) 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?

Six of the monitoring tools are considered to contribute to the environmental CAP specific objectives (SO4, SO5, SO6). The monitoring of agri-environmental and climate measures is explicitly mentioned for some tools. The implementation of area-based interventions seems to be easier to monitor with the tools identified in the inventory.

[66] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R0746

[67] https://publications.jrc.ec.europa.eu/repository/handle/JRC112918

[68] ECA Special Report 04/2020: Using new imaging technologies to monitor the Common Agricultural Policy: steady progress overall, but slower for climate and environment monitoring, 2020, doi:10.2865/46869

[69] https://wikis.ec.europa.eu/download/attachments/86968800/JRC127678_final.pdf?version=1&modificationDate=1682601334749&api=v2

[70] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021R2116&qid=1702232434387

A couple of tools require a special mention in this evaluation. The first one is the Italian Classyfarm system which bridges the communication gap between farmers and authorities, striving to elevate the safety and quality standards of products within the agri-food chain, that it contributes to Cap specific objective 1 monitoring. The other mention is for the Spanish Digital Farm Book that contributes to the cross-cutting objective by fostering the digitalisation of the agricultural sector.

On the other hand, many tools in the inventory are considered to contribute to all CAP specific objectives. Nevertheless, it is not clear how the data provided by these tools contribute to certain objectives. For instance, it is difficult to establish the contribution of the Area Monitoring System to the objective on jobs, growth and equality in rural areas.

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 these 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?

The IACS and LPIS has been in place during the previous programming period (only within the CAP, and not for other policies as suggested by our survey respondents). The novelty for this new programming period is the inclusion of the Area Monitoring System within the IACS, which is compulsory for all MS. Regulation 2116/2021[71] defines the AMS as "a procedure of regular and systematic observation, tracking and assessment of agricultural activities and practices on agricultural areas by Copernicus Sentinels satellite data or other data with at least equivalent value". This concept is what is behind of some of the tools identified in the inventory (iSIP, 3STR – Monitorig, for instance). A transitional period until the end of 2024 for the introduction of the AMS has been proposed during which MS are expected to obtain enough experience in the use of this tool.

Granularity

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

Evaluation Question: Can the tools support analysis and monitoring at different levels, including local, regional, national and EU level? What is the level of detail that these tools can reach? Do the tools provide results at different temporal resolutions (e.g. results expressed in days, months, years)?

The main purpose of the tools under this category is to verify that beneficiaries meet the requirements of the interventions they are benefiting from. The IACS collects information at beneficiary level and links the subsidy to the parcel by means of the LPIS. The georeferenced information provided by both the AMS and the geo-tagged photos are also linked to the subsidy by the LPIS. Therefore, the information collected by IACS presents a high level of granularity. In this direction, the Digital Farm Book gathers information at beneficiary level, providing a high level of granularity as well. As indicated by our survey respondents, the temporal resolution is minor than one year.

The information collected by these tools could be scaled-up at different spatial level. Regarding the temporal resolution, due to the fact that these systems are designed to verify the compliance of the beneficiary during the current season, they deliver data once a year.

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.

[71] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021R2116&qid=1702232434387

Ease of use

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

Evaluation Question: Do these tools require prior training or experience, or do they require much time to be able to use them properly? Can these tools be handled by a wide range of professionals, or do they require specific professional profiles to be implemented? Is there complex technology behind the use of these tools?

The most common tool for compliance monitoring identified in the inventory is the AMS. This system relies on new data acquisition technologies such as remote sensing images and geo-tagged photos. While remote sensing has been used for CAP control tasks (Check with Remote Sensing, CwRS), the AMS aims to fully exploit the capabilities of this data source. The use of this information indeed requires certain expertise on the acquisition, the data processing and the integration of the outputs into the IACS. On the other hand, geo-tagged photos have to be taken by ad-hoc IT applications, ensuring not only the inalterability of the geographical position of the picture but also the secure transmission to the public authority control system. Therefore, personnel with expertise in various fields such as remote sensing, ITC development and Policy knowledge are needed. Training on the meaning of the outputs of these systems is a key element as well. As suggested by our survey respondents, more than 100 hours training are generally needed.

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 their availability (e.g. costs, language, ease of use, available support, complexity, technical requirements)? Are there legal or patent-related limitations to their use?

Regulation 2116/2021 defines an Area Monitoring System as a procedure of regular and systematic observation, tracking and assessment of agricultural activities and practices on agricultural areas by Copernicus Sentinels satellite data or other data with at least equivalent value. Sentinel data are available through the Copernicus data space ecosystem[72]. Additionally, there have been other initiatives like Sen4CAP[73] and JRC-CbM for the CAP[74] aim to facilitate Sentinel data acquisition and processing by the public administrations in the context of CAP Checks by Monitoring. However, it is unclear from the inventory whether MSs have employed these systems or, on the contrary, they have developed their own procedures when implementing the AMS. Geo-tagged photos are considered data of equivalent value to Sentinel data. The inventory includes a couple of examples on smartphone applications developed by respective public administrations that are designed to send authenticated, geolocated photos of land plots.

Based on the inventory, it appears that each MS has developed its own ad-hoc tool, utilising the above-mentioned technologies to address country-specific needs. Consequently, it seems that each AMS is not accessible to other users.

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.

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.

[72] https://dataspace.copernicus.eu/

[73] http://esa-sen4cap.org/

[74] https://wikis.ec.europa.eu/display/GUIDANCEANDTOOLSFORCAP/Checks+by+monitoring

Evaluation Question: How much time does it take to set up this sub-category of tools (i.e time necessary to organise and prepare the implementation: e.g. installation of the software, get authorisation and credential to use it etc.)? And how much time does it take to implement the process of the tool (i.e. time necessary to implement the tool until the achievement of the intended outputs)?

The inventory does not inform on the time that took for MSs to implement the Area Monitoring System. We can only estimate a temporal frame based on the regulation to answer this question.

Previous to the entering into force the EU Regulation 2116/2021[75], the EC amended the Regulation 809/2014[76] in 2018 by Regulation 746/2018[77] which introduces the concept of Checks by Monitoring (CbM) as a verification method of beneficiaries' compliance with the legal bases of certain interventions. The CbM can be considered a first step towards the final implementation of AMS. According to the JRC[78], by 2018, there were only 5 public administrations adopting the CbM. In 2020, the European Court of Auditors issued a report[79] on "Using new imaging technologies to monitor the Common Agricultural Policy: steady progress overall, but slower for climate and environment monitoring" recommending a catalogue of best practices for CbM adoption. Based on the findings of the implementation of the CbM, in 2019, the EC introduced the concept of AMS as integral element of the IACS. Finally, in December 2021, the Regulation 2116/2021 establishes that "MSs shall set up and operate an area monitoring system, which shall be operational from 1 January 2023. If the full deployment of the system from that date is not feasible due to technical limitations, MSs may choose to set up and start the operation of such a system gradually, providing information for a limited number of interventions only. However, by 1 January 2024, an area monitoring system in all MSs shall be fully operational".

In view of the presented milestones, we estimate that the setting up of the AMS would have taken, in the shortest scenario, one year. In the case of those MSs that went for a gradual implementation of the tool, it would have taken three years, being the tool fully operational by time the Regulation 2116/2021 entered into force.

Cost-effectiveness

Cost-effectiveness refers to the extent to which the tool is able to achieve its outcomes within a reasonable amount of 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 budget can be necessary for the maintenance of the tool or subscription for patent-based tools every year?

The inventory does not provide with any insights into the budget invested in implementing the AMS. Indicatively, our survey respondents suggest that the set-up cost should be higher than 50,000 Euro, but without more specific indication. The main authors identified in the inventory are public administrations, Ministries of Agriculture, statistic institutions or entities dependant on certain public bodies. It is unclear whether these organisations developed the tool by themselves or they had to resort to private companies. However, we could consider the nature of investments these entities may have undertaken if they had developed the system by their own — hardware, software development (licensing) or, alternatively, cloud processing computing and specialised personnel.

Previous research[80] concludes that total annual administrative costs of IACS are estimated at between €1.7bn and €1.9bn. Figure 36 in this report presents that around 15% of the IACS related costs counts for IT investments and less than 5% is dedicated to horizontal IACS staff. The integration of the Area Monitoring System into the IACS is expected to increase the latter two budgetary items. Additionally, the quality assessment of the AMS (or the CbM) would also require budgetary reserves. The report concludes that "the average annual cost of IACS administration is estimated at €10 per hectare of Utilised Agricultural Area", spanning this cost across MS from 2 to 200 Euro and

[75] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021R2116&qid=1702232434387

[76] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014R0809

[77] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R0746

[78] https://joint-research-centre.ec.europa.eu/system/files/2019-04/17-ams_dgagri.pdf

[79] https://op.europa.eu/en/publication-detail/-/publication/4d5c725e-8ac0-11eb-b85c-01aa75ed71a1

[80] https://terraevita.edagricole.it/wp-content/uploads/sites/11/2019/07/costipac_final-report_en.pdf

recognise that small MSs have higher administrative cost per hectare of UAA (Utilised Agricultural Area). The extent to what these budgetary items will increase is expected to be highly dependent on the MS that is, size, regionalisation, technology adoption, flexibility of previous systems.

The study published by the EC in 2018 on the administrative costs of IACS[81] indicates that set-up costs (those that cover all costs arising from establishing or upgrading IACS to meet the legislative framework, typically one-off investment in IT, which is often outsourced) are estimated at €250 million to €275 million in the EU. These costs are higher at the beginning of the planning period and decrease over time. Cost estimates show that set-up costs also have a disproportionate impact on smaller MSs ranging from around 6% in larger MSs to up to 40% in smaller MSs. IT choices and architecture have a major impact on the variability of costs and the ability of MSs to adapt to changing regulations. Besides, the greater the level of customisation the greater the cost.

Human Resources Efficiency

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

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

Regarding the AMS, the professional skills required to properly implement the system encompasses a wide range of profiles. According to our survey respondents, these tools' implementation usually requires 50-250 people at least. The implementation of the AMS requires professional developers for developing new data acquisition and processing software. Additionally, integrating this new information into the IACS demand personnel with an ICT profile. The AMS is based on the regular and systematic observation, tracking and assessment of agricultural activities and practices on agricultural areas by Copernicus Sentinels satellite data or other data with at least equivalent value. To carry out these tasks not only expertise in remote sensing but also in geographical information management is needed. Finally, the judgment of the adequacy of beneficiaries to be eligible for CAP support call for personnel with knowledge of the legal bases of the requirements.

[81] https://op.europa.eu/en/publication-detail/-/publication/dabd45ab-9baf-11e9-9d01-01aa75ed71a1

3.4.3. Performance monitoring tools

Summary

***Accuracy**: moderate to high. These tools' accuracy varies due to different data sources and methods, facing challenges in integration and effectiveness across MSs.*

The data accuracy of performance monitoring tools is overall moderately high but varies based on the data sources and collection methods used. These tools already incorporate data collection protocols, validation and verification processes, quality checks, and cross-verification to ensure data accuracy. At the farm level, data is considered by the respondents of the survey as traceable and based on real observations rather than estimates or synthesis. Tools such as Decision Support Systems, which provide information for both farmer compliance and policy performance monitoring, face challenges in integrating into a systematic framework due to legal, organisational, and interoperability issues across MSs. Surveys targeting specific groups also offer quantifiable, comparable data with high traceability and transparency, particularly where traditional methods are inadequate. The EU's proposed Performance Monitoring Tools, managed by National Authorities, aim to centralise information for managing agricultural funding. However, their effectiveness is limited by inconsistent application across MSs and influenced by regional differences in agricultural practices, economic conditions, and environmental factors. While detail varies, all tools provide national-level results.

***Reliability**: high. These tools are considered highly reliable, based on concrete, well-documented and verified data, within an EU regulatory framework.*

The survey results indicate a strong consensus on the reliability of performance monitoring tools, primarily due to their foundation on concrete evidence sourced from well-documented and verified data. These tools, often managed by National Statistical Institutes or Ministries of Agriculture, adhere to stringent data collection protocols and methodologies defined under a regulatory framework set by Regulation (EU) 1306/2013, which outlines specific objectives for assessing the performance of the CAP. This framework facilitates external evaluation and review and ensures data and results are publicly accessible, transparent, consistent, and replicable. Additionally, the reliability of these tools is enhanced by the triangulation of multiple data sources, both at local and EU centralised levels, which contribute to a comprehensive data collection and verification process. However, there are challenges related to technical interoperability, especially when metadata or identifiers are not standardised, not allowing integration and comparison of data across different systems.

***Applicability**: moderate to high. These tools can contribute the monitoring of most CAP objectives, combining high-resolution monitoring with traditional data sources.*

Performance monitoring tools contribute to several SOs and phases in the EU policymaking process, particularly under the framework new CAP delivery model. According to the respondents to the survey, these tools contribute to the majority of SOs, and, in terms of policy stages, these results are useful, especially for enhancing financial allocation and CSP implementation monitoring. The survey results show that some MSs have used the tool in the previous CAP programming period. The trend in the EU is towards more automated, technology-driven monitoring, with many MSs adopting tools offering high spatial and temporal resolution. Data for assessing CAP's overall performance, typically collected from Ministries and National Agencies by entities like Eurostat, is aggregated from farm to regional and national levels, usually on an annual or bi/triennial basis. However, digital tools providing real-time data could offer a dynamic and timely perspective, complementing the more traditional, bureaucratically intensive data sources.

***Accessibility**: limited. The proposed tool is accessible only to specialised user groups and requires advanced training due to its complexity.*

The different performance monitoring tools results as complex, since are based on complex technologies involving GIS, remote sensing, data analytics, and advanced database integration across multiple software platforms. The survey highlights challenges in accessing relevant documentation or guidelines, and the necessity for specialised training and experience. Over 100 hours of training is deemed necessary for systems like the Common Monitoring and Evaluation System 2014-2020 and the SP CAP Performance Framework 2023-2027. These complexities limit

access to specific user groups and underscore the need to consider costs related to computing power and staff training. While "budget monitoring tools" are only accessible in certain MSs, the analysis of statistical data and CAP evaluation framework are available to stakeholders across all MSs.

***Efficiency**: limited. Setting up CAP performance monitoring tools is time-consuming and costly, requiring extensive data integration, compliance, and diverse expertise.*

The setting up of tools for CAP performance monitoring can be time-consuming, often taking several months to over a year due to extensive data collection, analysis capabilities, and integration with other systems. Legal compliance, data privacy considerations, and bureaucratic processes, along with the need for user familiarisation, further extend implementation time. While tools designed for real-time monitoring show immediate results, those for long-term analysis or policy impact assessment may take years to prove effective. Costs can be high, ranging from free open-source software to expensive proprietary systems requiring high-level hardware, servers, and increased computational power, leading to significant initial and ongoing technological investment costs. Maintenance costs include continuous technical support, and the implementation requires diverse professional expertise, including IT specialists, data analysts, legal and policy experts, and technical advisors, to handle technical setup, training, end-user support, and sector-specific analysis.

Analysis by criteria

ACCURACY

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

DATA ACCURACY AND VALIDITY

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

Evaluation Question: Are measures in place to avoid mistakes and redundancy and to ensure consistency in monitoring and/or the data collection activities (e.g. to verify the quality and reliability of the collected information)? Are collected data triangulated, corroborated, or substantiated by other data sources to check validity/accuracy? Are the collected data traceable back to its original source? To what extent is the data collected based on actual/direct observations rather than on estimations?

When asked whether the validity of the input data has been checked and verified, our survey respondents agreed by scoring (on average) 4.7 out of 5. Depending on the data sources, different assumptions on data accuracy can be made. IACS systems that rely on farm registry data incorporate information on parcels during beneficiaries' applications for CAP support (Reumaux et al., 2023), which are then cross-checked by other data sources such as remote sensing, area monitoring systems, and other technologies like geotagged photos. This integration, when feasible, simplifies the process but also improves the accuracy and reliability of the collected data offering increased validity and a quantified accuracy or uncertainty metric (e.g. "Traffic Lights System" to assess farmer applications).

Decision support systems that monitor in-situ agricultural variables are a source of information so that the collected data are even more traceable to its original source. Data confidentiality issues can arise, and there is an ongoing debate on whether those data could be used not only for monitoring purposes but for individual farmer compliance check (e.g. a way of "punishing" farmers based on data the digital decision support systems offer). In any case, the EU advice a clear separation between advice and checks by the MSs[82].

Mechanisms and methods for aggregating and sharing datasets from such systems (e.g. Farmer's Calendar, Farm Logs, Field books) for the needs of policy performance monitoring are still in various stages of incorporation into a systematic framework depending on the MS. In this context, recording of information in systems like the Field Book is mandatory in some cases, but still prone to fragmentation and errors, both intentional and unintentional. These are some of the main legal, organisational and interoperability challenges that remain in this sector. To mitigate

[82] https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A32013R1306

these inconsistencies, additional data sources such as farm machinery, geotagged photos, environmental sensors, and hard copies of invoices are used to corroborate the manually entered data.

Surveys that generate structured responses from the target groups (e.g. in Chech Republic)[83], providing quantifiable and comparable data for CSP monitoring and evaluation are also used to monitor policy performance where traditional methods might be inadequate, costly, or difficult to perform. The accuracy of those surveys depends on several factors that must be appropriately addressed by each authority conducting the survey as the data collected are used for various CAP output or result indicators calculation.

In some instances, such as in the Netherlands, data collection is entrusted to research institutes which have direct access to all relevant information, enhancing accuracy and traceability. Usually, national statistical authorities or other data collection agencies via the Ministry of Agriculture advisors in each country provide data collected to European statistical authorities (e.g. Eurostat) for publication via the Farm Accountancy Data Network FADN. In this case, different sampling thresholds across EU MSs are implemented to achieve a precise depiction of the agricultural sector in the EU. However, there have been concerns raised regarding the comparability of data and the potential underrepresentation of small farms.

Tool Representativeness

Representativeness refers to 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: Are the tools able to capture the diverse facets, variations, or elements within the system or dataset it seeks to represent? What is the potential geographical/sectorial/stakeholder coverage of these tools? Are these tools able to gather information from different sectors, stakeholders, and countries/regions?

Typically, the tools are operated by EU national authorities (e.g. Ministries of Agriculture) to address their specific needs by centralising and categorising the available information, while ensuring processing and administration of the European Agricultural funding at a national level. The tools are applicable to various crop indicators, farming practices and environmental conditions found across different EU regions. They can monitor and report various metrics on cultivation practices, crop types, soil health, environmental conditions, and agricultural inputs like fertilisers and pesticides, catering to different stakeholders needs in the agricultural sector, including farmers (when a decision support system is used), advisors, and paying agencies.

Their effectiveness in measuring performance is sometimes limited by variations in coverage across different MSs, sectors (e.g. economic data could be more accurate and representative than environmental data in a MSs or vice versa), and stakeholder groups (farmers, advisors, rural development organisations, agribusiness companies might have conflicting interests) but also differences within MSs that are attributed to regional disparities in agricultural practices, economic conditions, and local environmental factors.

Trying to deal with this phenomenon, targeted interventions with variable thresholds in performance indicators and (usually) annual targets are requested from the European Commission during approval process of the submitted CSPs[84].

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.

[83] https://www.szif.cz/en/about_us

[84] https://agriculture.ec.europa.eu/common-agricultural-policy/cap-overview/cap-2023-27_en

Replicability

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

Evaluation Question: Can the data collected by these tools be replicated successfully by others or through different approaches? Do the data remain consistent with other data collection activities or tools? Does collected data remain consistent over multiple data collection rounds (i.e. the collected information is always the same if data collection is replicated)?

Most tools are managed by national statistical institutes or ministries of agriculture with their unique approaches and methods. For larger MSs, there are subdivisions into regional organisations that conduct surveys or other data collection methods. A unified methodology that is applied and replicated on different levels of granularity can contribute to data consistency and replicability in this case.

In other cases, tools deal with challenges regarding technical interoperability, where data is in a machine-readable format (like CSV or JSON), and semantic interoperability, where the data's meaning is explicit and standardised across different systems using metadata and identifiers. This is due to the fact that multiple sources (local, national or EU centralised sources) and databases are involved in the process of data collection/assimilation.

Overall, while there is potential for consistency, especially introduced by tools mentioned in Checks by Monitoring (CbM)[85], more comprehensive frameworks are needed to improve the adaptation and comparability of methods and data especially in the social and financial sectors and not so much in technical tools that rely on remote sensing data. Finally, it is a fact that the EC had to remind MSs to provide self-assessments, indicating that while the proposed system is designed for stability, its effectiveness can vary depending on the implementation by MSs.

The survey does not offer specific insights into aspects of data replicability, or data consistency within or across different tools. However, it is important to mention that here is an high level of agreement among respondents on the fact that the data sources are clearly identified and documented, the validity of data has been checked and verified (e.g. through standardised data collection protocols, data quality audits), the tool is based on well-founded evidence and is considered reliable (degree to which the results can be reproduced by repeating the process of implementation of the tool under the same conditions).

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?

According to our survey, all respondents agree on the transparency and traceability of the data sources, and that input data is clearly identified and documented (scored 5 out of 5 on average). However, the availability of such information can vary based on the specific tool, the managing authority, and the regulations governing data sharing and transparency within the EU framework. Regulation (EU) No 1306/2013[86] specifies the objectives against which these tools measure the performance of the CAP, suggesting that there is a framework for external evaluation and review. In most cases, data and results are publicly available and capable of describing the social, economic, and environmental conditions of the country and the changes that occur in it. Most methodologies are generally not easy to find for external analysis and in some cases, they are not publicly accessible. MSs are responsible for performing the actual assessment of the inspection procedures under CAP. All relevant documentation of this process, including ground truth collection and validation, should be available for information and later verification (such as audits). This

[85] https://wikis.ec.europa.eu/display/GUIDANCEANDTOOLSFORCAP/Checks+by+monitoring

[86] EUR-LEX - 32013R1306 - EN - EUR-LEX. https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:32013R1306

implies that there is a level of transparency and accessibility of externals in the methodologies and results used in these processes[87].

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? Have these tools (i.e. their collected data) 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?

The reviewed tools are designed to address various policy objectives and steps in the policymaking process particularly in the context of the CAP post-2020 reforms related to specific CSPs of each MS. According to our survey respondents, the identified tools can contribute to all CAP several specific objectives.

In general, every MS establishes a link between result indicators described in their CSPs and broader CAP specific objectives[88]. This direct link is measured via quantitative targets and milestones using a set of result indicators that support the new CAP delivery model. Also, innovative actions with Agricultural Knowledge and Innovation Systems impact the creation or alteration of national/regional CAP policies, as indicated by the new CAP through co-creating innovative agro-environmental-climate interventions and eco-schemes[89]. Although the relationship between individual inputs and outputs, and between overall CAP results and impacts, at regional and at national level, is usually not straightforward, contribution to the CAP specific objectives and the CSPs design and monitoring in made in varying degrees depending on the MS.

Article 121(4) of Regulation 2115/2021[90] states that the annual performance reports must include specific information about realised outputs, realised expenditure, realised results and distance to respective intervention targets, in a way to control and monitor performance effectively.

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 these 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?

The agricultural sector differs significantly between the MSs, and current EU rules give MSs certain guidelines, defining how to apply the CAP. In May 2017, the EU paying agencies signed the "Malta Declaration," encouraging the use of new technologies for simplifying IACS (Integrated Administration and Control System). The first agency in Italy started using this new approach in 2018. By 2019, 15 paying agencies across Belgium, Denmark, Italy, Malta,

[87] https://wikis.ec.europa.eu/display/GUIDANCEANDTOOLSFORCAP/Checks+by+monitoring

[88] https://agriculture.ec.europa.eu/system/files/2023-10/pmef-cover-note-indicators_en.pdf

[89]

[90] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32021R2115

and Spain had adopted 'checks by monitoring' for some of their schemes[91]. Furthermore, out of 66 paying agencies, 15 used Copernicus Sentinel data in 2019 to check aid applications for various schemes and beneficiary groups. This use of new technologies indicates a growing trend towards more automated and technology-driven monitoring processes within the EU as dictated by the new CAP. For implementing other CAP policies, several databases are used by the EC and MSs. These include the Information System for Agricultural Market Management and Monitoring (ISAMM), the Clearance Audit Trail System (CATS), and the Information System for Agriculture Refund Expenditure (AGREX). Monthly declarations of expenditure are transmitted by MSs, and the EC publishes an annual report on the implementation of direct payments, providing metrics regarding the financial aspects of CAP measures[92].

Granularity

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

Evaluation Question: Can the tools support analysis and monitoring at different levels, including local, regional, national and EU level? What is the level of detail that these tools can reach? Do the tools provide results at different temporal resolutions (e.g. results expressed in days, months, years)?

All tools that have high spatial (e.g. NUTS 3) and temporal (e.g monthly) resolution incorporate farm digital field books and farm registries. These data are usually referring to farm-level or regional level after data aggregation methods. To aggregate into a larger region (e.g. NUTS 1) or at national level the inclusion of other agricultural information sources is the most common approach especially when these data are used to calculate indicators and performance reports (usually conducted at an annual basis depending on the respective indicators), mentioned in the CSPs.

Most of impact and context indicators that used to assess the overall CAP policy performance are collected from ministries and national statistics agencies by Eurostat, Joint Research Centre, European Environment Agency, etc., on annual, two- or three-years delay. This data can indeed provide a useful progress measure towards targets established by not only CSPs, but the overall performance review of the EU objectives. Issues with the excessive number of indicators and sub-indicators, which can complicate obtaining a clear picture of CAP achievements have been observed too, and since they are subject to delays, they negatively impact their usefulness for “early” monitoring[93].

In contrast, data derived from farm calendars, satellite sources, and other digital tools offer more 'near real time' data, usually available at the end of the growing season or at even higher frequencies[94]. These technologies in most cases “complement” the existing (bureaucratically intensive) statistical and administrative data sources, providing a more dynamic and timely insight into CAP's implementation and performance.

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: Do these tools require prior training or experience, or do they require much time to be able to use them properly? Can these tools be handled by a wide range of professionals, or do they require

91

92 https://mef4cap.eu/storage/files/MEF4CAP_D3.1_Review_of_current_monitoring_systems%20_1_.pdf

93 https://mef4cap.eu/storage/files/MEF4CAP_D3.1_Review_of_current_monitoring_systems%20_1_.pdf

94 https://dataspace.copernicus.eu/news/2023-10-19-cap-monitoring-national-scale-slovenia-based-copernicus-data

specific professional profiles to be implemented? Is there complex technology behind the use of these tools?

The tools reviewed, such as the IACS in Greece, Italy's New Monitoring System AGEA, and the Czech Republic's data collection methodologies, are based on complex technologies involving geographic information systems (GIS), remote sensing, data analytics, and advanced database integration and multiple software platforms. Based on our information, we were unable to locate specific official documentation or user guidelines provided by the entities responsible for non-digital tools providing details on the level of education, specialty, experience, and training that is required by professionals to possess in order to use the tools. According to our survey respondents, implementing these tools require more than 100 hours of training.

However, digital tools provide more information. For example, AgriSnap tool[95] for geotagging photos is an easy-to-use application that allows users to take photos, add notes, and submit them to the respective national authorities. Training is provided to farmers, advisors, and administrative staff to ensure that they can use the tool effectively. Also, tools that are part of Checks by monitoring do provide documentation and technical guidance with a database of markers suitable to work with the IACS and the LPIS. Technical knowledge of Earth Observation theory and geodatabase management are essential to variable selection like the number and size of parcels or cloudiness of the area, that lead in selecting the type of images and deciding on the IT infrastructure size and host premises.

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 their availability (e.g. costs, language, ease of use, available support, complexity, technical requirements)? Are there legal or patent-related limitations to their use?

Many monitoring tools are designed for use by government agencies, research institutions, and professionals in agriculture and environmental management and are restricted to specific user groups with high complexity and technical requirements. They do require significant investment in software, hardware, and training that influences negative the ease of adoption and use by other MSs or organisations. Also, based on our research tools developed in one EU country are primarily available in that country's language, limiting their broader use or accessibility. However, the "Checks by Monitoring" (CbM)[96] documentation published by the EC offers comprehensive technical guidance, methodologies, reports, on open-source tools that mainly utilise digital technologies and satellite data such as the use of geotagged photographs, geodata and technologies for agriculture. These resources are designed to assist MSs' organisations in addressing the previously mentioned challenges in a more interoperable and accessible framework with standardised practices across MSs.

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 sub-category of tools (i.e time necessary to organise and prepare the implementation: e.g. installation of the software, get authorisation and credential to use it etc.)? And how much time does it take to implement the process of the tool (i.e. time necessary to implement the tool until the achievement of the intended outputs)?

[95] https://play.google.com/store/apps/details?id=ie.gov.agriculture.geotag.app&hl=el&gl=US) for geotagging photos is

[96] https://wikis.ec.europa.eu/display/GUIDANCEANDTOOLSFORCAP/Checks+by+monitoring

Simpler tools that target a broad set of users (e.g. farmers) require only a few minutes to download and install e.g. a smartphone device with internet access[97]. More complex digital tools require days to a few weeks to set up and implement by skilled and authorised personnel, especially if they come with comprehensive official specifications, guidelines, and documentation. More complex systems, (e.g. Sen4CAP in a CREODIAS virtual machine was tested in Greece)[98] involving extensive data collection, analysis capabilities, qualitative surveys with questionnaires, integration with other systems, could take several months, up to a year or even more. Robust workflows for setting up data migration, testing, troubleshooting, and communication protocols must also be designed, implemented, and tested.

Users often need time to become familiar with those tools, which usually includes training sessions and trial/pilot runs. Legal compliance checks and data privacy agreements along with bureaucratic hurdles can also significantly increase the time necessary to implement the tool until the achievement of the intended outputs. Tools designed for "near" real-time data monitoring might show results immediately, whereas a tool used for long-term data analysis or policy impact assessment could take years to demonstrate its effectiveness.

Cost-effectiveness

Cost-effectiveness refers to the extent to which the tool is able to achieve its outcomes within a reasonable amount of 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 budget can be necessary for the maintenance of the tool or subscription for patent-based tools every year?

Software licensing or purchase costs of acquiring related software can range from free (open-source) to thousands of Euro for proprietary systems that support high-level hardware requirements such as servers, specialised equipment and increased computational power which can add to the initial and recurrent technological investment costs. A more specific example is described in Greece[99], where considering the hardware requirements, software needs, and personnel efforts for a medium to low size MS, Sen4CAP tool total costs can easily overcome 6,000 to 12,000 Euro annually. These estimated costs include a virtual machine with 92GB RAM, 8 Cores, and subsequent upgrade to 8TB of storage, alongside the software engineering efforts for installation, troubleshooting, and upgrades, with the overall financial commitment to escalate further with various user experience research and feedback collection costs.

Continuous technical support and occasional troubleshooting can add to the maintenance costs that usually burden the organisation that owns or implements the tools. Annual or monthly fees apply to patent-based solutions while national organisations with state-owned infrastructure usually subcontract regular (e.g. every year) updates or upgrades that are necessary, to comply with evolving regulatory requirements and technological advancements. For example, setting up and implementing a tool like Sen4CAP seems to be better suited for installation on a DIAS (Data and Image Access Services) (that means recurrent monthly costs) rather than a local server, unless significant initial investment in resources like a dedicated IT engineer and server with administrative rights are available [1].

Human Resources Efficiency

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

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

Professional expertise required for the implementation of a tool and the analysis of its results depend on several factors, including the complexity of the tool, the scope of its use, and the specific objectives aiming to achieve in the

[97] https://play.google.com/store/apps/details?id=ie.gov.agriculture.geotag.app&hl=el&gl=US

[98] https://www.niva4cap.eu/wp-content/uploads/2021/09/D3.5Recommendations-for-standardised-connections-between-IACS-and-other-applications_v1.0.pdf

[99] https://www.niva4cap.eu/wp-content/uploads/2021/09/D3.5Recommendations-for-standardised-connections-between-IACS-and-other-applications_v1.0.pdf

context of CAP performance monitoring. Essential roles include: IT specialists, system administrators and developers that can handle the technical aspects of designing, installing, setting up, integrating, configuring, and maintaining the software, staff for training and support to end-users after tool implementation, specialists and advisors with sector-specific knowledge (e.g., agriculture, environment, finance), data analysts, legal experts and policymakers/

Team size can vary depending on the implementation stages and the lifecycle of the tool. For less complex tools, a small team of individuals may be sufficient but in large organisations or for tools requiring significant customisation, the team could include dozens of individuals across different institutions (regional, national or even at EU level) and roles.

3.4.4. Data and knowledge stocktaking tools

Summary

***Accuracy**: moderate to high. The tools provide essential farm-based data for the CAP, sourced from various data points, and validated through appropriate measures, standardised collection, and quality checks.*

Different are the tools employed for gathering and managing data necessary for policy monitoring and evaluation, as well as offering insights into the state of farm-based CAP regulation. This includes FADN and IACS, together with other Statistics, such as desk research, Eurostat, survey data, and national data. These tools rely on multiple data sources and according to 90% of survey respondents this, together with other measure in place, such as, data validation and verification processes, standardised data collection protocols and quality audits, helps in avoiding redundancy. These tools capture data with different spatial and temporal granularity, mainly at the national level, and are most effective for national-level analysis, followed by EU and regional levels, but with a frequency of updates annual or multiannual, not allowing proper close to real-time information. However, limitations exist, such as the inability of current data and tools to deliver certain critical elements required for well-informed policymaking. There is a notable lack of automated data collection processes at the farm level and the development of IACS and FADN varies among different countries, impacting the consistency and efficacy of data collection and analysis.

***Reliability**: high. Systems like FADN and IACS, often validated and updated, use overlapping data sources, with most survey respondents finding their methodologies transparent and accessible.*

Many data and knowledge stocktaking tools are based on common data sources and some of the data collected by these tools can be replicated in other tools. According to 70% of survey respondents, the results from the tools are validated and regularly updated with the latest information. The most common validation methods are comparison with other monitoring tools/methods or cross-validation from different data sources to ensure consistency. For FADN, sources like farm invoices, accounting offices, veterinarians, suppliers, and banks are particularly valuable. Conversely, the IACS system is managed by Paying Agencies. The survey results show that only 11% of respondents reported that the tools are not accessible to all end-users. Generally, methodologies and results are transparent and widely available.

***Applicability**: high. These tools support various agricultural policy steps and objectives, with widespread use across MSs and applications extending to other areas like environment and climate policies and monitoring.*

Tools under this sub-category, like IACS and FADN systems, contribute to various SOs and policy steps in agriculture, with IACS aiding in agricultural subsidy parameters and animal-based CAP interventions and FADN focusing on economic data. They are particularly useful in designing CSP indicators, needs assessment, SWOT analysis, intervention design, monitoring CSP implementation, and ex-ante analysis. According to the results of the survey, the data and knowledge stocktaking tools have been used across all MSs in the previous programming period, with 80% of respondents reporting their usage and 40% also declaring their application in other policy areas like environment and climate policy.

***Accessibility**: moderate. Data and knowledge stocktaking tools shows a significant variation in accessibility, mainly due to the different IT systems in place at EU level.*

The use and accessibility of data and knowledge stocktaking tools vary significantly across the EU, due to the differing structures and complexities of the IT systems implemented in various MSs (MS). This is particularly evident in some cases, where some MS have complex and challenging systems, while others maintain very basic setups. Due to these disparities, it is difficult to define a uniform standard for usability, accessibility, and complexity at the EU level. Nonetheless, in a recent survey, these tools received an overall score of 4.2 out of 5, indicating a generally high level of user satisfaction. However, 50% of the participants noted variability in the availability of these tools. Among the data collection systems evaluated, EUROSTAT recorded the highest score, signifying its strong reputation in terms of usability and effectiveness.

***Efficiency**: moderate. MSs' variations significatively affect the implementation complexity, costs, and setup times of data and knowledge stocktaking tools, with diverse human resource needs and wide cost ranges.*

The differences among MSs significantly influence the complexity and costs of implementing and managing the proposed data and knowledge stocktaking tools, especially during changes. The required human input varies and needs to be specialised. In the case of FADN, the total number of people needed to provide data varies greatly by MS, heavily influenced by the sample size. For IACS, it is important to note that some use in-house development teams, while others outsource part or all development to third-party companies. This variation depends on the levels of digitalisation and administration setup, which also affects the complexity and costs of managing the tools, particularly when modifications occur. Overall, the costs are very high and vary depending on the specific tool and the country of implementation. Most respondents to the survey reported costs ranging from less than 1,000 Euro to over 50,000 Euro for the same tool. The setup time also varies significantly by tool type, with about 50% of respondents indicating that more than 6 months are needed. However, the required time can differ, but a year for the first installation, with subsequent updates, modifications, and training.

Analysis by criteria

ACCURACY

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

DATA ACCURACY AND VALIDITY

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

Evaluation Question: Are measures in place to avoid mistakes and redundancy and to ensure consistency in monitoring and/or the data collection activities (e.g. to verify the quality and reliability of the collected information)? Are collected data triangulated, corroborated, or substantiated by other data sources to check validity/accuracy? Are the collected data traceable back to its original source? To what extent is the data collected based on actual/direct observations rather than on estimations?

Many monitoring and data collection tools for data and knowledge stocktaking facilitate the collection, storage and management of data and information necessary for policy monitoring and evaluation, informing policy choices. Some of these tools borne for other purposes, for example many are monitoring tools, but they also act as a data pool. Examples are FADN systems[100] which collects annually information in the MSs for each of the 80,000 FADN sample farms which includes approximately 1,000 variables and refers to physical and structural data (location, crop areas, livestock numbers, etc.); and economic and financial data (production value of the different crops, stocks, sales, purchases, production costs, assets, liabilities, production quotas and subsidies, etc). Also, IACS system seeks to efficiently categorise farmer applications based on a set of predefined factors, ensuring a streamlined process for application verification and subsequent actions. Additionally, it aids in collecting valuable data for understanding various dimensions of agricultural activity. IACS is fundamental in consolidating various data sources to make informed decisions regarding payments. It ensures all relevant data is evaluated before making payment decisions. Official statistics also offer insights into the current state of farms based on the CAP regulation, utilising various data sources, including desk research, Eurostat, FADN data, survey data, and national data By analysing this data, stakeholders can assess the economic condition and compliance levels of the farms under scrutiny.

About 90% of our survey respondents declare that measures are in place to avoid mistakes and redundancy in the input data. Most often, these measures consist of data validation and verification processes (7 respondents), standardised data collection protocols (6 respondents) as well as data quality audits (4 respondents).

The fact that most of these tools use and interlink other data sources, redundancy are in many ways avoided and it guarantees the accuracy and validity of data. Furthermore IACS (which is based also on the AMS and LPIS) collects data based on direct (satellite and geophoto) observations. FADN has a basic interlink with IACS (for example some of the data of Italian FADN comes from IACS and ISTAT). In Netherland the FADN system collect data from several sources.

[100] https://agriculture.ec.europa.eu/data-and-analysis/farm-structures-and-economics/fadn_en

For the IACS, the validity and accuracy of data is well explained above. The information delivered by the tools under this sub-category is combined to make a judgement on the beneficiaries' compliance. This combination could be seen as a sort of cross-check system to ensure the accuracy and validity of the data. Therefore, the information provided by these tools is accurate enough for its purpose. For instance, Earth Observation and geo-tagged photos are integral parts of the Area Monitoring System (AMS). The AMS enables the assessment of the requisites of some area-based interventions. Thus, the AMS, along with the Land Parcel Information System (LPIS), are two of the building blocks of the Integrated Administration and Control System (IACS). The information stored in the IACS may be combined with the information coming from different administrative databases and registries within the so-called Traffic Light Systems. The Traffic Light Systems deliver a judgment on the compliance (green light), non-compliance (red light), or non-conclusive (yellow light) state of each potential beneficiary applying for subsidies.

Tool Representativeness

Representativeness refers to 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: Are the tools able to capture the diverse facets, variations, or elements within the system or dataset it seeks to represent? What is the potential geographical/sectorial/stakeholder coverage of these tools? Are these tools able to gather information from different sectors, stakeholders, and countries/regions?

The respondents to our survey indicate that this sub-category of tools provided results at different spatial levels (multi-national, national, regional). However, when asked about the spatial resolution of the tool i.e. the capacity to observe data at different spatial level the respondents confirm that despite offering the opportunity to investigate different levels, their potential is mostly in analysis at national level, followed by EU level and Regional level. Most of our survey respondents (70%) report that the results of the tools are validated. The validation methods adopted in the majority of cases is either comparison with other existing monitoring tools/methods or cross-validation from different data sources, indicating that checks are made to assess consistency with other tools.

Generally speaking the tools envisaged in the subcategory can capture the variations and the geographical and stakeholder coverage they want to represent, gathering information from different sectors. Systems as FADN, IACS, and Offical Statistics collect data coming from all the MSs covering different elements. As mentioned in the section above, IACS collects valuable data for understanding various dimensions of agricultural activity to parameter the subsidies to agriculture, FADN focused on economic and financial data, official statistics leaded by the National Statistical Institutes (NSIs) collect and store micro (farm level) data.

Furthermore other tools as Digital Farm Book collect all the activities implemented at farm level. This module is starting to be integrated in IACS system. Especially for IACS and FADN, in the case of regionalise countries (French, Iatly, etc) they can gather information from different regions.

Besides, it is also true that the Commission holds large amounts of data for CAP design, monitoring and evaluation[101]. The Commission uses conventional tools such as spreadsheets to analyse the data it collects from the MSs. Current data and tools do not deliver certain significant elements (e.g. details of the environmental practices applied, and off-farm income) that are needed for well-informed policy-making. The Commission has taken several legislative and other initiatives to make better use of existing data, but barriers to making the best use of collected data remain.

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.

[101] https://op.europa.eu/webpub/eca/special-reports/agri-big-data-16-2022/en/

Replicability

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

Evaluation Question: Can the data collected by these tools be replicated successfully by others or through different approaches? Do the data remain consistent with other data collection activities or tools? Does collected data remain consistent over multiple data collection rounds (i.e. the collected information is always the same if data collection is replicated)?

Regarding replicability, systems like IACS can produce consistent results repeating the data collection twice. Other systems like FADN can be less consistent, but in any case, since the systems are very different from country to country, it is very difficult to assess precisely. However, data collection systems are very diverse across MSs, and their degree of replicability might be variable.

Some of the data collected by these tools can be replicated in other tools. An example is the Italian FADN, which uses data from IACS, from national statistics and from national databases on the number of heads or livestock units. Or in Netherland where the collection of data for the Dutch FADN relies on a variety of sources (with systematic bookkeeping serving as a common starting point for both economic and environmental data). Farm invoices are a particularly valuable source of information, offering insights into not just the cost, but also the type and quantity of inputs used, such as pesticides. Other sources are: accounting offices, veterinarians, banks, and suppliers.

Regarding IACS, a database system is set up and operated in each MS to administer and control direct payments and some rural development payments, such as payments under an agri-environment scheme. Article 68 of Regulation 1306/2013[102] stipulates that an IACS must include: a computerised database; an identification system for agricultural land (the Land Parcel Identification System – LPIS - system). Another requirement is the mandatory use of computer-aided geographical information techniques (GIS), which also includes air photographs or satellite pictures; a system for the identification and registration of payment entitlements; an integrated control system; a system of recording the identity of each farmer who submits an aid application or payment claim; and a system for the identification and registration of animals if the MS receives relevant payments. IACS is implemented by the Paying Agency. This has developed into a collection of different systems between MSs, even between regions within MSs. As a result, IACS has 44 different implementations across MSs[103]. Implementation model depends on the national organisation and relations between farmers, farmer organisations, paying agencies, control agencies and other involved parties. The information contained in the IACSs is therefore also different, although the common base consists of the required recording of specific attributes related to the regulation. The IACS system can also have a broader use, e.g. incorporating national schemes. The environment in which the IACS systems evolve are quite complex. IACS systems operate under national IT infrastructures and architectures, including data. Therefore, not all IACS datasets are homogeneous. Also, systems are developed in house or by external contractors. Another complexity is the way the IACS systems interact with farmers. All MSs have introduced digital portals for farmers to view and edit data, but the level of automatic linkage of data to farmer's own systems is diverse. In particular, the new data requirements for proof supporting the claims, such as geotagged photos or farm machine data, cannot be entered manually.

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?

When asked about the availability and accessibility of methodologies for all end-users, the majority of respondents to our survey (56%) declares to not know, followed by a (33%) who consider the methodologies available and accessible. Only 11% declared that the methodologies are not available. The majority of respondents (70%) report

[102] https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A32013R1306

[103] See www.niva4cap.eu.

that the tool is regularly updated to incorporate the latest available information. A 40% of respondents also declare that the tool allows feedback from end-users, followed by 30% not knowing and a 30% affirming that the feedback is not allowed.

Regarding FADN, in contrast to the availability of standard results, access to farm-level data, a feature that obviously adds greatly to its value as a research tool, is universally restricted, respecting the general principle of maintaining confidentiality. On IACS side, the system is composed of various data components and tools (for instance, the tools developed by the NIVA project[104]), and dissemination of underlying methodologies can be more compelx. The IACS system is managed by Paying Agencies and has strong links with farmers who provide data through the GSAA component and who received data about controls and payments. The IACS system receives input from external data providers, for instance orthoimages and data on regulated areas. The IACS system is also exporting data towards EU systems and other users.

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? Have these tools (i.e. their collected data) 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?

Some of the tools envisaged in this subcategory can address manifold policy objectives. According to our survey respondents, this sub-category of tools contributes to all CAP Specific objectives, and can be used especially to support activities of design of CSP Indicators, needs assessment and SWOT analysis, ex-ante analysis, design of interventions, and of course for monitoring the implementation of the CSPs. However, since these tools rely on complex systems, they cannot be changed (e.g. re-adapted to other purposes) easily. Also, new regulations might be required along with technical modifications.

For example, IACS can manage, monitor and serve for MSs to control all the area and animal-based common agricultural policy (CAP) interventions (such as direct payments interventions and area and animal-based rural development interventions) and ensures that comprehensive and comparable data is available throughout the EU[105]. Monitoring agricultural, environment and climate policies helps to keep track of the impact of the CAP, its environmental performance and progress towards EU targets. It relies on a range of data sources. These include meteorological data and forecasts, existing maps and statistics, positional information and remotely sensed data. IACS is responsible for Pillar I and II of the CAP.

Regarding the FADN, this system is not only the one which provides annually key data and information that allows assessing and evaluating the economic and financial performance of all EU farms, but it can also be used for: i) evaluating production costs; ii) helping the public decision maker to support the planning and definition of interventions, the implementation of programs and measures, the monitoring and evaluation of what is financed and the implementation of "services" for the benefit of beneficiaries; iii) implementing thematic analyses; iv) designing the PSP (Economic justification of public support for Direct Payments (basic and redistributive) and for some Eco-

[104] https://www.niva4cap.eu/uploads/D.3/D3.4%20Recommendations%20for%20IACS%20%20data%20%20flows%20v1.0%20w.pdf

[105] https://agriculture.ec.europa.eu/common-agricultural-policy/financing-cap/assurance-and-audit/managing-payments_en

schemes, Economic justification of support for Payments Coupled with plant production (durum wheat, sunflower, rice, etc.) and animal production (dairy cows, sheep and goats, etc.), Economic justification for some agro-climatic-environmental interventions (ACA) of Rural Development (integrated production, organic production, animal welfare, etc.), Economic justification for companies located in disadvantaged areas, Calculation of the economic size of the agricultural company (in terms of Standard Production) for access to some measures of the Rural Development Program (PSR), Selection of beneficiaries and/or operations based on the criteria identified in the programs (e.g. specific production specialisations).

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 these 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?

Some of the tools (as IACS and FADN) are mandatory, and all the MSs must use them[106,107]. It means that all the MSs has a long lasting experience in using these tools, which have been used in the previous programming periods. Regarding the use of these tools for other policies, IACS and FADN started to look at environmental and climate policies, IACS with the use of eco-schemes (as a future CAP innovation, eco-schemes shall provide support for farmers who observe agricultural practices beneficial for the environment and climate. It is a measure to reward and incentivise farmers for taking action towards a more sustainable farm and land management with the objective to maintain public goods. The participation for farmers is voluntary. MSs are however obliged to include one or more eco-schemes in their CSPs) and FADN with its future transformation into FSDN which will foresee the collection of environmental and social variables.

Granularity

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

Evaluation Question: Can the tools support analysis and monitoring at different levels, including local, regional, national and EU level? What is the level of detail that these tools can reach? Do the tools provide results at different temporal resolutions (e.g. results expressed in days, months, years)?

The respondents to our survey indicate that this sub-category of tools can provide results at different spatial levels (from the farm to EU level). However, often their focus in terms of results remains mostly at national level. Regarding the temporal resolutions, some tools as FADN or IACS provide data on yearly bases, where Digital Farm Book can provide data on a daily or monthly bases.

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: Do these tools require prior training or experience, or do they require much time to be able to use them properly? Can these tools be handled by a wide range of professionals, or do they require specific professional profiles to be implemented? Is there complex technology behind the use of these tools?

[106] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2021.435.01.0187.01.ENG

[107] https://eur-lex.europa.eu/eli/reg_del/2022/1172/oj

The results of our survey would suggest an overall satisfaction of respondents regarding the ease of use of these tools (scored on average 4.2 out of 5). However, very different tools are employed across the EU, likely with varying degree of ease of use that is difficult to assess.

The tools envisaged under the subcategory present a great heterogeneity between the MSs. Infact, MSs are often responsible for designing their own Information Management (IM) approach and design the IT systems and tools to support this. The Commission does not interfere with national decisions or governance. It means that is difficult to say whether all of them are easy to use or not. Talking about the FADN we know that in some MSs the systems are quite complex and not easy to use, and in other they are very basic (Vrolijk and Poppe, 2019).

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 their availability (e.g. costs, language, ease of use, available support, complexity, technical requirements)? Are there legal or patent-related limitations to their use?

Regarding the tools under these categories, and especially IACS and FADN, they can vary from MS to MS. It means that we can find tools easy to use and not complex, and other more complex and difficult to use. It not easy to define a common situation. In fact, when asked about the availability of the tools, 50% of our survey respondents declared that they can be available, but the remaining share of respondents either answered they are not available (30%) or they don't know (20%). However, in general, there is a vast array of technologies, developed during the past years across all MSs, that can allow the set up of this sub-category of tools. It is important to consider also that some of these tools (e.g. those related to IACS and FADN) are already in place in all MSs, as they are mandatory.

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. 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 sub-category of tools (i.e time necessary to organise and prepare the implementation: e.g. installation of the software, get authorisation and credential to use it etc.)? And how much time does it take to implement the process of the tool (i.e. time necessary to implement the tool until the achievement of the intended outputs)?

IT solution as IACS and FADN are quite complex systems (most of the time) which need a reasonable amount of time for their implementation. In some countries the adoption of an IACS system passes by a public tender which identify an external provider of the system. Normally are quite long lasting process. One year for the first installation and then updating, modification and trainings. For the FADN systems, which in many cases has been developed by the IT department of the liaising agencies, it is a bit more complex to estimate the time for their development. It is a long term development process which crosses different steps and improvements of the technical solutions (e.g. learning by doing). In line with the analysis of compliance and performance monitoring tools (which often are the same tools), we can conclude that time needed can vary, but generally, long periods are required, especially for the set-up.

Cost-effectiveness

Cost-effectiveness refers to the extent to which the tool is able to achieve its outcomes within a reasonable amount of 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 budget can be necessary for the maintenance of the tool or subscription for patent-based tools every year?

As for performance and compliance monitoring tools, the cost of setting up and implementing these tools can be very high.

FADN: very little information on the costs of FADN. A study conducted in 2015 revealed a great difference between MSs[108]. For the voice data collection by Liaison Agency the cost varies from 182,000 Euro in Malta to 6,972,304 Euro in Belgium. For the voice data collection by public advisory services the cost vary from 569,182 Euro in Latvia to more than 18 million Euro in Poland.

For IACS[109,110], It consists of different databases and so allows farmers to enter an online aid application and administrations to manage the applications and claims. The higher costs related to non-IACS payments is related to the requirements for project-based measures under Rural Development. IACS running costs are estimated at 12 % of overall IACS administrative costs, while the bulk of costs (74 %) relates to administration of claims and controls. With the 2013 reform, overall burden for national administrations is estimated to have increased by one third. Factors behind this increase include the introduction of greening provisions, the added heterogeneity from increased tailoring of policies and the modernisation of IT systems and digitalisation of controls. With regard to beneficiaries, the administrative burden is estimated at around 6 % of public expenditure.

In 2017, the average annual costs of IACS administration per country was around 65 million Euro. There is, however, significant variation across MSs and years, ranging from estimated annual costs of 2 million to more than 300 million Euro. Per hectare of utilised agricultural area (UAA), the annual IACS cost of administration is estimated at 10.47 Euro. Across the EU, the annual IACS cost ranges from 2 to 208 Euro per hectare of UAA, indicating considerable variation across MSs. Similarly, average annual IACS costs per agricultural holding are estimated at 168.61 Euro. The data shows large differences between smaller and larger MSs, with a disproportionately high cost for smaller MSs. Administrative tasks linked to the IACS are more burdensome for smaller MSs that do not benefit from economies of scale. For example, all MSs need to implement an IACS IT infrastructure but, although the absolute cost of IACS IT infrastructure is higher for large MSs, the cost per beneficiary is lower.

Human Resources Efficiency

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

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

As we defined above, many of these tools differ significantly across MSs in terms of technological and technical components and, therefore, in terms of necessary resources. According to our survey respondents, the implementation of the tools can vary depending on the specific tool, some 50% of respondents indicate less than 5 people can be necessary, whereas others suggested between 5 to 10 people are necessary (e.g. for using Bledjacam in Malta), and between 25 and 50 persons are needed for data collection in Poland.

Regarding IACS[111], some use in-house development teams and others outsource part or all the development to third party companies. Further, some MSs have a network of third-party groups or structures on hand to assist or receive digital applications and transfers, while others provide web-enabled services that allow beneficiaries to enter their

[108] https://op.europa.eu/en/publication-detail/-/publication/02ee48a9-d479-11e5-a4b5-01aa75ed71a1

[109] https://op.europa.eu/en/publication-detail/-/publication/4a92d9e8-d7f9-49c9-af3c-444acff6ec04

[110] https://op.europa.eu/en/publication-detail/-/publication/598b81ff-dfbc-11e7-9749-01aa75ed71a1

[111] https://op.europa.eu/en/publication-detail/-/publication/dabd45ab-9baf-11e9-9d01-01aa75ed71a1

applications directly. And, in addition, levels of digital sophistication of beneficiaries and administrations also differs across and within MSs. Differences in MSs choices over the various elements outlined above have the potential to influence the complexity and costs of implementing and managing IACS, particularly when changes take place.

Regarding FADN[112], the total number of FTEs (full time equivalents) required to provide data to FADN varies enormously by MS, although this is heavily influenced by the sample size. The lowest number of FTEs required (not including MSs for which not all FTEs are known) is Sweden at 5.0, followed by Latvia and Cyprus with 8.0 FTEs each. The highest number of FTEs is required in Poland (375.3), followed by Italy (97.1) and the UK (83.8).

[112] https://op.europa.eu/en/publication-detail/-/publication/02ee48a9-d479-11e5-a4b5-01aa75ed71a1

3.4.5. Key findings

In this section, we summarise and cast light on the key aspects emerging from the evaluation of monitoring and data collection tools, and we conclude by identifying the strengths and weaknesses of this category of tools.

Monitoring and data collection tools have a significant role in the context of the CAP to enhance the effectiveness and efficiency of policy implementation and evaluation processes. These tools include CAP compliance monitoring tools, which ensure that beneficiaries meet specific intervention requirements and criteria, and performance monitoring tools, which provide insights into MSs' progress towards their objectives by assessing result indicators. Similarly, data and knowledge stocktaking tools are also relevant in policy monitoring and informed decision-making since they are employed in collecting, storing, and managing data processes.

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

Accuracy is assessed as moderate-to-high (performance monitoring, data collection tools) or high (compliance monitoring toold). Overall, evaluating monitoring and data collection tools indicates high data accuracy, offering the ability to capture different spatial and temporal granularities from various data points. These are usually delivered yearly and refer to the current season, validated through appropriate quality checks and cross-checked using different data sources. Nevertheless, the main limitation in accuracy can be related to the various data collection methods employed at the country level, as well as the notable lack of automated and unified data collection methods at the farm level and interoperability among the different tools at the EU level, as in the case of IACS and FADN.

Reliability is rated as high for the three sub-categories of tools. Data triangulation and data validation procedures ensure these tools' reliability. Overall, the survey results indicate that the tools are accessible to most EU level end-users, providing a good level of transparency. For compliance and performance monitoring tools, a regulatory framework is established by Regulation (EU) 1306/2013, outlining specific objectives for assessing the performance of the CAP. Some limitations, such as sampling and data aggregation issues, can partially hinder the reliability of some data systems for data stocktaking tools.

Applicability of the tools is assessed as moderate (compliance monitoring tools), moderate-to-high (performance monitoring tools) or high (data collection tools). They support several agricultural policy steps and CAP objectives, with widespread use across MSs. Most of them were used in the previous programming period and are considered effective in measuring compliance with environmentally related measures, demonstrating a high level of granularity. This is particularly true for performance monitoring tools, which can provide spatial data at various levels, ranging from parcel-level information to geotargeted data. According to survey respondents, these tools are also valuable for various policy steps, particularly in CAP implementation monitoring. Additionally, the past application of monitoring and data collection tools to environmental and climate policies highlights the relevance of these tools for various policy areas.

Accessibility is limited (performance monitoring tools), limited-to-moderate (compliance monitoring tools) or moderate (data collection tools). This limitation is partly due to the need for access by users with various levels of expertise and specialised groups. These users require specific training to operate these complex systems effectively. Such complexities restrict access to these tools to particular user groups and highlight the importance of considering the costs associated with computing power and staff training. Additionally, there is a lack of uniformity in the IT systems used at the EU level, as most MS develop their own applications tailored to their specific needs. This is also the case with Area Monitoring Systems (AMS), which is employed for CAP monitoring purposes. Starting from 1 January 2024, due to Regulation EU 1306/2013, these systems began to be operational at the EU level.

Efficiency is limited (performance monitoring tools), limited-to-moderate (compliance monitoring tools) or moderate (data collection tools). In most cases, the time required to set up the system is substantial, often taking several months to over a year. Additionally, acquiring the appropriate skills and the need for user familiarisation further extend the implementation time. The costs for implementing monitoring and data collection tools vary depending on the complexity of the systems among the different MS, but overall, they are considered high. This is particularly evident in the case of data and knowledge stocktaking tools and certain AMS, where it has been reported in the survey that the costs could require an investment of more than 50,000 Euro.

In conclusion, monitoring and data collection tools are highly accurate and reliable, with moderate to high applicability across various policy areas. However, challenges include diverse data collection methods at the country level, limited accessibility due to complex system requirements and training needs, and overall limited efficiency marked by significant setup time and high implementation costs. Despite these challenges, the tools are integral to effective policy implementation and offer valuable insights with the potential to be applied in environmental and climate policy monitoring. While current Regulations make it mandatory for MSs to set up some monitoring tools, MSs are also taking initiatives to develop new monitoring systems. Yet, this development process is likely to take a few years, and the new tools that will be set up will be relevant for monitoring the next-generation CSPs.

4. Conclusions and next steps

The Tools4CAP project aims to take stock of the tools implemented so far by MSs to support the design and monitoring of the CSPs to identify, leverage and promote the replication of the most promising tools and good practices. To this end, Tools4CAP delivered an online inventory mapping the tools used by MSs for the different tasks of policy design and monitoring. The objective of this report (Deliverable 1.3) is to evaluate the identified tools for their accuracy, reliability, applicability, accessibility and efficiency to cast light on the pros and cons of different tools and provide a reference for MSs to understand the implications of different types of tools. Based on our results, benchmarking factsheets were produced to highlight the strengths and weaknesses of various tools to support different policy tasks.

Our evaluation covered about 80 tools included in the inventory and other tools used by the EC, grouped into four main categories of tools, further broken down into a total of 10 sub-categories of tools based on their methodological characteristics. The evaluation relied on the Tune-up Evaluation Framework (TEF), consisting of five criteria, 12 sub-criteria, and corresponding 12 evaluation questions. Different information sources were considered to answer the evaluation questions, notably interviews, online surveys, national focus groups, desk research and consortium expert knowledge.

Overall, the results show that stakeholder needs assessment and policy choice supporting tools tend to be less accurate and reliable but more applicable, accessible and efficient, mainly due to the relatively low costs, knowledge required and methodological complexity. On the other hand, policy analysis and monitoring tools, which are technically more complex and often rely on a technological component, can show much higher accuracy and reliability but limitations related to applicability, accessibility, and efficiency.

Stakeholder needs assessment tools, encompassing online consultation and survey tools, workshops and conferences, and focus groups and meetings, exhibit varying degrees of accuracy, reliability, applicability, accessibility, and efficiency. While online consultations, surveys, workshops, and conferences demonstrate limited accuracy due to sampling issues and reliance on stakeholders' perceptions, focus groups and meetings offer slightly higher accuracy owing to expert knowledge and methodological protocols. Reliability is deemed moderate for all tools, with online consultations, surveys, and focus groups more replicable than workshops and conferences. Despite concerns about accuracy and replicability, these tools are highly applicable and accessible, with the flexibility to adapt to different contexts and policy steps. They are widely available and user-friendly, though workshops and conferences may require more preparation time and incur higher costs. While open consultation can be successfully combined with more targeted approaches based on focus groups or workshops, the choice of the stakeholder needs assessment tools depends heavily on the policy step, which is intended to be informed by the stakeholders' perspectives. These tools can be supplemented with policy analysis tools to address accuracy and reliability limitations, particularly for technical tasks like intervention or target setting.

Policy choice supporting tools encompass two sub-categories, each with its nuances. Voting and prioritisation tools exhibit limited to high accuracy, with oversimplification and strategic behaviours posing challenges, while data-driven approaches offer potential improvements. Multicriteria analysis tools, relying on precise data, demonstrate moderate to high accuracy, prioritising transparency and validation processes, although capturing information from diverse governance levels remains an unaddressed concern. Reliability is generally rated as limited or moderate, with voting tools facing complexity issues and multicriteria suffering from inconsistency due to methodological variations. Applicability is considered moderate to high, with potential for tailoring to various objectives and policy steps, albeit uncertainties persist regarding governance levels. Accessibility is assessed as moderate to high, with varying levels of skill requirements across tools. Efficiency ranges from moderate to high, influenced by design complexity and resource allocation, making these tools valuable for complex policy decisions despite their limitations. Complementing these tools with stakeholder needs assessment and policy analysis tools can enhance accuracy and reliability, fostering coherent consensus in policy decision-making processes.

Policy analysis tools, crucial for evidence-based decisions, demonstrate moderate to high accuracy and reliability, with statistical analysis tools leveraging reliable data sources and descriptive statistics, while simulation models tools ensure traceability and consistency through data verification. Statistical analysis tools offer moderate to high reliability, thanks to their standardised nature, although varying methods may yield different outcomes, while simulation models tools exhibit reliability but may generate divergent results for the same scenarios. Applicability is moderate to high for both tools, with statistical analysis tools widely used for data processing and analysis and simulation models tools assessing multiple policy objectives. However, accessibility varies, with statistical analysis tools being widely available and easier to use, while simulation models tools require specific expertise and training.

Efficiency is generally limited to moderate for both tools, with statistical analysis tools being moderately efficient. Still, simulation models require significant time and resources to set up and implement due to complexity. Despite their importance, policy analysis tools face limitations in accessibility, costs, and skills, underscoring the need for investments and peer learning among MSs, and centralised support by the EC for the implementation of more complex simulation models.

Monitoring and data collection tools demonstrate moderate to high accuracy and reliability, underpinned by high data accuracy and data triangulation procedures, ensuring robustness in capturing various spatial and temporal granularities. However, challenges persist in the diversity of data collection methods at the country level and the lack of unified data collection methods at the EU level, impacting accuracy. Applicability is rated as moderate to high, supporting multiple agricultural policy steps and objectives across the MSs, particularly effective in measuring compliance with environmentally related measures. Nevertheless, limited accessibility due to complex system requirements and training needs poses barriers, while implementation costs are generally high, ranging from substantial setup times to significant financial investments. Despite these challenges, these tools remain integral to effective policy implementation, although the development process for new monitoring systems may take several years.

A first important consideration is that, based on the evaluation results, we cannot refer to “best” tools, rather we can identify most “proper” tools. This is because any tool has pros and cons, and its utility depends on the specific country's needs. Consequently, a tool might be irrelevant in one country but crucial in another.

A second important consideration is that different tools can accomplish different tasks, and because MSs are assigned several design and monitoring tasks, it is for MSs to choose the proper mix of tools. This exercise underscores the imperative of adopting an integrated strategic approach to leverage the strengths and approaches of each tool complementarily. Therefore, the decision to use one tool over another should optimally be made based on a comprehensive assessment of tool availability and future CAP requirements. Most often, it is not just about preferring a tool over another, but it is about finding the appropriate combination of tools to carry out the different policy tasks in a coherent and meaningful way within the time and resource constraints. This underscores the necessity to make explicit and facilitate information flux between the various tools to promote transparency and evidence-based decision-making. Timely utilisation of these tools is critical to optimise the information flux, enabling swift responses to emerging challenges and opportunities. Consequently, a cohesive strategy integrating the targeted use of different tools in a timely manner should be developed, aligning their deployment with overarching objectives and operational timelines to enhance the overall effectiveness of CAP design. In this regard, Tools4CAP will provide guidance to the selection of tools by producing a roadmap for the uptake of tools and methods (Deliverable 3.3) and a handbook of good practices (Deliverable 5.3).

A third consideration regards the need to explore other tools available in science or used for other policies. While MSs proposed a wide range of tools for certain categories (e.g. stakeholder needs assessment and monitoring tools), fewer MSs have set up policy choice supporting and policy analysis tools. This evidence is captured by our online inventory of tools. Plenty of analytical tools are already available that could be adapted to the CAP context. This potential needs to be further exploited.

The results of this evaluation underpin future work, and in particular, the identification of the most promising tools to be re-adapted (Work Packages 2, 3 and 4) and tested in a number of case studies (Tools4CAP Work Package 5). Also based on the evaluation results, a list of potentially interesting tools will be identified by the consortium, and scrutinised by MSs’ authorities, which will express their interest and select a number of tools to be showcased in at least 10 case studies. In addition, the TEF will be re-adapted into a self-assessment instrument for MSs to verify the coherence and consistency of their approach to the design and monitoring of the CSPs.

The report, as well as the benchmarking factsheets, can help MSs gain a comprehensive view of the range of tools potentially useful for next-generation CSPs and to grasp the advantages and challenges to account for when implementing different tools. As such, this report can serve as a reference manual for end users.

5. References

Abatan, S. M., Olayemi, M. (2014). The role of statistical software in data analysis. International Journal of Applied Research and Studies (iJARS), ISSN 2278-9480. https://www.researchgate.net/publication/360335782_The_Role_of_Statistical_Software_in_Data_Analysis

Abel, E., Keane, J., Paton, N.W., Fernandes, A.A.A., Koehler, M., Konstantinou, N., Cortes Rios, J.C., Azuan, N.A., Embury, S.M. (2018). User driven multi-criteria source selection. Information Sciences, 430–431, 179-199. https://doi.org/10.1016/j.ins.2017.11.019.

Alif, Ž., Šumrada, T., Žgajnar, J. (2023). Does economic situation cause land abandonment? Estimating economic viability of farming in a Sub-Mediterranean region in Slovenia. In DROBNE, S. (Ed.), SOR '23: Proceedings of the 17th International Symposium on Operational Research in Slovenia (pp. 11-14). Ljubljana: Slovenian Society Informatika - Section for Operational Research.

Anggraeni, Mustika, Gupta, Joyeeta, Verrest, Hebe J.L.M. (2019). Cost and value of stakeholders participation: A systematic literature review. Environmental Science & Policy, 101, 364–373. https://doi.org/10.1016/j.envsci.2019.07.012.

Arsham, H. (2020). Excel for statistical data analysis. https://www.researchgate.net/publication/344638517_Excel_For_Statistical_Data_Analysis

Bedrač, M. et al. (2023). Prva ocean stanja v kmetijstvu. https://www.kis.si/Porocila_o_stanju_v_kmetijstvu/

Bergen, N., Labonté, R. (2020). “Everything is perfect, and we have no problems”: Detecting and limiting social desirability bias in qualitative research. Qualitative Health Research, 30(5), 783-792.

Berkhout, J., Crepaz, M., Hanegraaff, M., Junk, W. M. (2023). Online focus groups as a tool to study policy professionals. Research & Politics, 10(4). https://doi.org/10.1177/20531680231211398.

Bertolozzi-Caredio, D., Severini, S., Pierre, G., Zinnanti, C., Rustom, R., Santoni, E. and Bubbico, A., (2023), Risks and vulnerabilities in the EU food supply chain, Publications Office of the European Union, Luxembourg, 2023, doi:10.2760/171825, JRC135290. https://publications.jrc.ec.europa.eu/repository/handle/JRC135290

Better Regulation Toolbox (2023). https://commission.europa.eu/law/law-making-process/planning-and-proposing-law/better-regulation/better-regulation-guidelines-and-toolbox_en

Bleich, E., Pekkanen, R. (2015). Transparency about qualitative evidence. Newsletter of the American Political Science Association Organized Section for Qualitative and Multi-Method Research, 13(1), 8-13. https://www.researchgate.net/publication/282340447_Data_Access_Research_Transparency_and_Interviews_The_Interview_Methods_Appendix

Bogonos, M., Blanco, M., Ciaian, P., Deppermann, A., Frank, S., Havlik, P., Jongeneel, R., Kremmydas, D., Perez-Dominguez, I., Petsakos, A., Tabeau, A., Witzke, H.P. (2020). D.3.1: Inter-model baseline harmonization and comparison. https://www.suprema-project.eu/images/Deliverable_D3.1_updated.pdf.

Brečko, J., Žgajnar, J. (2022). Possible impact of risk management strategies with farm model on a mixed farm type. Business Systems Research, 13(3), 23-35.

Brečko, J., Žgajnar, J. (2023). Price volatility and its impact on farm operations, an example of analysis with a farm model. In: DROBNE, Samo (ed.), et al. SOR '23: Proceedings of the 17th International Symposium on Operational Research in Slovenia: Bled, Slovenia, September 20-22, 2023 (pp. 19-22). Ljubljana: Slovenian Society Informatika - Section for Operational Research.

Brenna, J., Hamilton-Jones, N., Kennedy, B., O’Connell-Hussey, E., Ramsay, A. (2021). “Evaluating the Common Agricultural Policy 2023-2027”. IGEES Internship, 93. https://assets.gov.ie/212485/3478c5fe-31a5-49f1-946d-4375e447f91d.pdf

Busetto, L., Wick, W., Gumbinger, C. (2020). How to use and assess qualitative research methods. Neurological Research and Practice, 2, 1-10.

Cagliero, R., Bellini, F., Marcatto, F., Novelli, S., Monteleone, A., Mazzocchi, G. (2021). Prioritising CAP Intervention Needs: An Improved Cumulative Voting Approach. Sustainability, 13(3997). https://doi.org/10.3390/su13073997

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. Italian Review of Agricultural Economics, 77(3), 25-40. DOI: 10.36253/rea-13717

Chaves-Avila, R., Gallego-Bono, J. R. (2020). Transformative policies for the social and solidarity economy: The new generation of public policies fostering the social economy in order to achieve sustainable development goals. The European and Spanish cases. Sustainability, 12(10), 4059.

Cinelli, M., Kadziński, M., Gonzalez, M., Słowiński, R. (2020). How to support the application of multiple criteria decision analysis? Let us start with a comprehensive taxonomy. Omega, Volume 96. https://doi.org/10.1016/j.omega.2020.102261.

Cyr, J. (2016). The pitfalls and promise of focus groups as a data collection method. Sociological Methods & Research, 45(2), 231–259.

Dace, E., Blumberga, D. (2016). How do 28 European Union MSs perform in agricultural greenhouse gas emissions? It depends on what we look at: Application of the multi-criteria analysis. Ecological Indicators, 71, 352-358.

Dean, M. (2020a). Multi-criteria analysis. In Niek Mouter (Ed.), Advances in Transport Policy and Planning (Vol. 6, pp. 165-224). Academic Press. https://doi.org/10.1016/bs.atpp.2020.07.001.

Dean, M. (2022b). Including multiple perspectives in participatory multi-criteria analysis: A framework for investigation. Evaluation, 28(4), 505-539. https://doi.org/10.1177/13563890221123822

Dehe, B., Bamford, D. (2015). Development, test and comparison of two Multiple Criteria Decision Analysis (MCDA) models: A case of healthcare infrastructure location. Expert Systems with Applications, 42(19), 6717-6727. https://doi.org/10.1016/j.eswa.2015.04.059.

Dyckhoff, H., Souren, R. (2022). Integrating multiple criteria decision analysis and production theory for performance evaluation: Framework and review. European Journal of Operational Research, 297(3), 795-816. https://doi.org/10.1016/j.ejor.2021.05.046.Economic and Social Council of Republic Bulgaria (2018) “Mountainous and semi-mountainous regions in Bulgaria - problems and opportunities for development", ESC/3/050/2018 (BG).

El Benni, N., Grovermann, C., Finger, R. (2023). Towards more evidence-based agricultural and food policies. Q Open, 3(3), qoad003. https://doi.org/10.1093/qopen/qoad003

Erjavec, E., Lovec, M., Juvančič, L., Šumrada, T., Rac, I. (2018). The CAP Strategic Plans beyond 2020: Assessing the architecture and governance issues in order to achieve the EU-wide objectives (p. 52). Study requested by the AGRI Committee, European Parliament, Policy Department for Structural and Cohesion Policies. Brussels, Belgium.

European Commission. (2020). Strategy on agricultural statistics Final version for publication. Retrieved from https://ec.europa.eu/eurostat/documents/749240/749310/Strategy+on+agricultural+statistics+Final+version+for+publication.pdf/9c7787ca-0e00-f676-7a64-7f56e74ec813

European Commission. (2021). Commission staff working document: Analysis of links between CAP reform and Green Deal (COM(2021) 815 final). Retrieved from https://www.europarl.europa.eu/RegData/docs_autres_institutions/commission_europeenne/com/2021/0815/COM_COM(2021)0815_EN.pdf

European Commission. (2023). Modelling Inventory and Knowledge Management System of the European Commission (MIDAS). Retrieved from https://web.jrc.ec.europa.eu/policy-model-inventory/explore/.

European Court of Auditors. (2022). 'Data in the Common Agricultural Policy: Unrealistic potential of big data for policy evaluation'. Special report, 58. Retrieved from https://www.eca.europa.eu/Lists/ECADocuments/SR22_16/SR_Big_Data_in_CAP_EN.pdf

Fern, E. F. (1982). The use of focus groups for idea generation: The effects of group size, acquaintanceship, and moderator on response quantity and quality. Journal of Marketing Research, 19(1), 1–13.

Fraussen, Bert, Albareda, Adrià, Braun, Caelesta. (2020). Conceptualizing consultation approaches: Identifying combinations of consultation tools and analyzing their implications for stakeholder diversity. Policy Sciences, 53(3), 473–493. DOI: 10.1007/s11077-020-09382-3.

Fraussen, Bert. (2023). Consultation Tools and Agenda-Setting. In Michael Howlett (Ed.), The Routledge Handbook of Policy Tools (pp. 149–159). Taylor & Francis.

Gammie, E., Hamilton, S., Gilchrist, V. (2017). Focus group discussions. In Hoque, Z., Parker, L. D., Covaleski, M. A. (Eds.), The Routledge Companion to Qualitative Accounting Research Methods (pp. 372–386). Taylor & Francis.

Gell, T. (2023). How Much Does a Focus Group Cost? Retrieved from https://www.driveresearch.com/market-research-company-blog/how-much-does-a-focus-group-cost-focus-groups-syracuse-ny/.

Gilliland, D., Melfi, V. (2010). "A Note on Confidence Interval Estimation and Margin of Error". Journal of Statistics Education, Volume 18, Number 1. Retrieved from www.amstat.org/publications/jse/v18n1/gilliland.pdf

Gongora-Salazar, P., Rocks, S., Fahr, P., Rivero-Arias, O., Tsiachristas, A. (2023). The Use of Multicriteria Decision Analysis to Support Decision Making in Healthcare: An Updated Systematic Literature Review. Value in Health, 26(5), 780-790. https://doi.org/10.1016/j.jval.2022.11.007.

Gonzalez-Martinez, A.R., Jongeneel, R., Salamon, P., Zezza, A., De Maria, F., Potori, N. (2021). The COVID-19 pandemic and the EU agri-food sector: MS impacts and recovery pathways. Studies in Agricultural Economics, 123, 153-158. https://studies.hu/the-covid-19-pandemic-and-the-eu-agri-food-sector-member-state-impacts-and-recovery-pathways/

Hair, J. F., Celsi, M., Ortinau, D. J., Bush, R. P. (2008). Essentials of Marketing Research. McGraw-Hill/Higher Education. McGraw Hill 6th Edition.

Hennink, M., Kaiser, B. N. (2022). Sample sizes for saturation in qualitative research: A systematic review of empirical tests. Social Science & Medicine, 292, 114523.

Jastram, Sarah M., Berberyan, Zara. (2023). Democratic Stakeholder Representativeness. Journal of Public Administration Research and Theory, Volume 33, Issue 4, October 2023, Pages 647–660. https://doi.org/10.1093/jopart/muad005

Jongeneel, R., Gonzalez-Martinez, A., van Leeuwen, M., Verhoog, D. (2022). Potential medium-run impacts of the Russia-Ukrainian war on the Dutch agriculture and food system: An assessment. (Report / Wageningen Economic Research; No. 2022-053). Wageningen Economic Research. https://doi.org/10.18174/571299

Simkus, J., (2023). Between-Subjects Vs. Within-Subjects Study Design, July 31, 2023. Retrieved from https://www.simplypsychology.org/between-subjects-vs-within-subjects-design.html

Kamali, F.P., Borges, J.A.R., Meuwissen, M.P.M., de Boer, I.J.M., Oude Lansink, A.G.J.M. (2017). Sustainability assessment of agricultural systems: The validity of expert opinion and robustness of a multi-criteria analysis. Agricultural Systems, Volume 157. https://doi.org/10.1016/j.agsy.2017.07.013.

Keating, Colma. (2003). Facilitation Toolkit: A practical guide for working more effectively with people and groups. https://library.dbca.wa.gov.au/static/FullTextFiles/001674.pdf

Killen, P. (2007). “Replication Statistics”. In Book Best Practices of Quantitative Methods (pp. 103-124). Retrieved from https://www.researchgate.net/publication/220034823_Replication_statistics

Kitzinger, J. (1995). Qualitative research: Introducing focus groups. British Medical Journal, 311(7000), 299–302.

Koop, G. (2013). “Analysis of Economic Data”. Wiley Publisher. E-book, 272.

Liamputtong, P. (2011). Focus Group Methodology: Principles and Practice. SAGE Publications Ltd. https://doi.org/10.4135/9781473957657

Lobe, B., Morgan, D., Hoffman, K. A. (2020). Qualitative data collection in an era of social distancing. International Journal of Qualitative Methods, 19, 1609406920937875.

Lochmiller, C. R. (2021). Conducting thematic analysis with qualitative data. The Qualitative Report, 26(6), 2029-2044.Lovrić, N., Fraccaroli, C., & Bozzano, M. (2023). A future EU overall strategy for agriculture and forest genetic resources management: Finding consensus through policymakers’ participation. Futures, 151, 103179.

Maier, J. F., Eckert, C. M., Clarkson, P. J. (2016). Model granularity and related concepts. Proceedings of International Design Conference - DESIGN 2016. Dubrovnik - Croatia, 1327–1336.

Manos, B., Bournaris, T., Chatzinikolaou, P. (2011). Impact assessment of CAP policies on social sustainability in rural areas: An application in Northern Greece. Operational Research International Journal, 11, 77–92. https://doi.org/10.1007/s12351-010-0078-y

Mardani, A., Jusoh, A., Nor, M. D. K., Khalifah, Z., Zakwan, N., Valipour, A. (2015). Multiple criteria decision-making techniques and their applications – A review of the literature from 2000 to 2014. Economic Research-Ekonomska Istraživanja, 28:1, 516-571. https://doi.org/10.1080/1331677X.2015.1075139

Matthews, A. (2021). Evaluating the Legislative Basis for the New CAP Strategic Plans.

Mazzocchi, G., Cagliero, R., Monteleone, A., Angeli, S., Tarangioli, S. (CREA PB) (2021). La prioritizzazione delle esigenze nel piano strategico nazionale PAC 2023-2027.

Ministry of Agriculture of the Republic of Latvia, (2023). Reģionālajos semināros visā Latvijā informē par atbalsta maksājumiem lauksaimniekiem un lauku uzņēmējiem 2023. gadā - SEMINĀRU NORISES GRAFIKS. Retrieved from https://www.zm.gov.lv/lv/jaunums/regionalajos-seminaros-visa-latvija-informe-par-atbalsta-maksajumiem-lauksaimniekiem-un-lauku-uznemejiem-2023-gada-seminaru-norises-grafiks

ML (Niedersächsisches Ministerium für Bundes- und Europaangelegenheiten und regionale Entwicklung), (2021). Online-Veranstaltung zum ELER-Förderkonzept 2023 - 2027. Zoom-Konferenz am 24. und 25.02.2021. Dokumentation der Veranstaltung. Available online at www.ml.niedersachsen.de%2Fdownload%2F167573%2FDokumentation_der_WiSo-Partner_Veranstaltung_am_24._und_25._Februar_2021.pdf

Munda, G. (2004). “Social multi-criteria evaluation (SMCE)”: Methodological foundations and operational consequences. European Journal of Operational Research, 158(3), 662-677. http://dx.doi.org/10.1016/S0377-2217(03)00369-2

Munda, G. (2008). Social Multi-Criteria Evaluation for a Sustainable Economy. Springer, Heidelberg, New York. ISBN: 978-3-540-73702-5. http://www.springer.com/business/operations+research/book/978-3-540-73702-5

Munda, G., Albrecht, D., Becker, W., Havari, E., Listorti, G., Ostlaender, N., Paruolo, P., Saisana, M. (2020). The Use of Quantitative Methods in the Policy Cycle. Science for Policy Handbook, Elsevier. https://doi.org/10.1016/B978-0-12-822596-7.00018-8.

Munda, G. (2021). Qualitative reasoning or quantitative aggregation rules for impact assessment of policy options? A multiple criteria framework. Quality & Quantity. https://doi.org/10.1007/s11135-021-01267-8

Murgado-Armenteros, E. M., Torres-Ruiz, F. J., Vega-Zamora, M. (2012). Differences between online and face to face focus groups, viewed through two approaches. Journal of Theoretical and Applied Electronic Commerce Research, 7(2), 73–86.

OECD. (2008). Understanding Economic Statistics: An OECD Perspective. OECD Publishing, Paris. https://doi.org/10.1787/9789264046986-en.

Peck, R., Olsen, C., & Devore, J. (2008). Introduction to statistics and data analysis (3rd ed.). Thomson. Pp. 885. Retrieved from https://www.spps.org/cms/lib/MN01910242/Centricity/Domain/859/Statistics%20Textbook.pdf

Pečnik, Ž., Žgajnar, J. (2022). Resilience of dairy farms measured through production plan adjustments = Odpornost kmetij s prirejo mleka z različnimi prilagoditvami proizvodnega načrta. Journal of Central European Agriculture: JCEA, 23(1), 207-219.

Perez Dominguez, I., Fellmann, T., Weiss, F., Witzke, H., Barreiro Hurle, J., Himics, M., Jansson, T., Salputra, G., & Leip, A. (2016). An economic assessment of GHG mitigation policy options for EU agriculture (EcAMPA 2). EUR 27973 EN. Luxembourg: Publications Office of the European Union. JRC101396. Retrieved from https://publications.jrc.ec.europa.eu/repository/handle/JRC101396

Philippidis, G., Helming, J., Tabeau, A. (2017). Model linkage between CAPRI and MAGNET: An exploratory assessment. EUR 28625 EN. Luxembourg: Publications Office of the European Union. ISBN 978-92-79-69101-0, doi:10.2760/712139, JRC106595.

Pisano, U., Lange, L. K., Lepuschitz, K., & Berger, G. (2015). The role of stakeholder participation in European sustainable development policies and strategies. Retrieved from https://www.researchgate.net/profile/umberto-pisano/publication/312496614_pisano_u_lk_lange_k_lepuschitz_and_g_berger_2015_the_role_of_stakeholder_participation_in_european_sustainable_development_policies_and_strategies

Prescott, T. L., Phillips II, G., DuBois, L. Z., Bull, S. S., Mustanski, B., & Ybarra, M. L. (2016). Reaching adolescent gay, bisexual, and queer men online: Development and refinement of a national recruitment strategy. Journal of Medical Internet Research, 18(8), Article e200.

Reumaux, R., Chopin, P., Bergkvist, G., Watson, C., & Öborn, I. (2023). Land Parcel Identification System (LPIS) data allows identification of crop sequence patterns and diversity in organic and conventional farming systems. European Journal of Agronomy, 149, 126916. https://doi.org/10.1016/j.eja.2023.126916

Richard, B., Sivo, S. A., Orlowski, M., Ford, R. C., Murphy, J., Boote, D. N., & Witta, E. L. (2021). Qualitative Research via Focus Groups: Will Going Online Affect the Diversity of Your Findings? Cornell Hospitality Quarterly, 62(1), 32-45. https://doi.org/10.1177/1938965520967769

Sahni, Deepak. (2007). A Controlled Experiment on Analytical Hierarchy Process and Cumulative Voting - Investigating Time, Scalability, Accuracy, Ease of use and Ease of learning.

Shaffer, D., & Serlin, R. (2004). What good are statistics that don't generalize? Educational Research, 33(9), 14-25. Retrieved from https://www.researchgate.net/publication/242547346_What_Good_are_Statistics_that_Don't_Generalize

Siciliano, G. (2009). Social multicriteria evaluation of farming practices in the presence of soil degradation: A case study in Southern Tuscany, Italy. Environmental Development and Sustainability, 11, 1107–1133. https://doi.org/10.1007/s10668-008-9169-9

Stewart, D. W., & Shamdasani, P. N. (1998). Focus group research: Exploration and discovery. In Bickman, L., & Rog, D. J. (Eds.), Handbook of Applied Social Research Methods (pp. 505–526). SAGE.

Tufail, H., Qasim, I., Masood, M. F., Tanvir, S., & Butt, W. H. (2019). Towards the selection of Optimum Requirements Prioritization Technique: A Comparative Analysis. In Proceedings of the 5th International Conference on Information Management (ICIM 2019), Cambridge, UK, 24-27 March 2019, IEEE Digital Library. pp. 227-231. DOI: 10.1109/INFOMAN.2019.8714709

Velasquez, M., & Hester, P.T. (2013). An Analysis of Multi-Criteria Decision-Making Methods. International Journal of Operations Research, 10(2).

Vogels, D. M., Howcroft, E., Tourtouri, E., & Demberg, V. (2020). How speakers adapt object descriptions to listeners under load. Language, Cognition and Neuroscience, 35(1), 78-92. DOI: 10.1080/23273798.2019.1648839

Volkov, A., Balezentis, T., Morkunas, M., & Streimikiene, D. (2019). Who benefits from CAP? The way the direct payments system impacts socioeconomic.

Vrolijk, H. C. J., & Poppe, K. J. (2019). Costs and benefits of collecting farm data for the new CAP's data needs: Empirical evidence. Wageningen University and Research, The Netherlands.

Wagenmakers, E., Sarafoglou, A., Aarts, S., Albers, C., Algermissen, J., Bahník, Š., van Dongen, N., Hoekstra, R., Moreau, D., van Ravenzwaaij, D., Sluga, A., Stanke, F., Tendeiro, J., & Aczel, B. (2021). Seven steps toward more transparency in statistical practice. Nature Human Behavior.

Žgajnar, J., Juvančič, L., Erjavec, E., & Kavčič, S. (2019). Scenarijska analiza ukrepov SKP po letu 2021: Dohodkovni učinki na ravni tipičnih kmetijskih gospodarstev Slovenije: Raziskovalna podpora za načrtovanje kmetijske politike po letu 2020. Ljubljana: Biotehniška fakulteta.

Žgajnar, J., Juvančič, L., Kavčič, S., & Erjavec, E. (2021). CAP post 2022 scenarios and income impacts: A case analysis for selected typical farms in Slovenia. Acta Agriculturae Slovenica, 117(2), 1-12.

Žgajnar, J., Kavčič, S., & Erjavec, E. (2021). Scenarijska ocena vpliva sprememb SKP po 2023 na ekonomski položaj tipičnih kmetijskih gospodarstev - rezultati končnega scenarija: Končno poročilo. Ljubljana: Univerza v Ljubljani, Biotehniška fakulteta.

Žgajnar, J., Kavčič, S., Tomšič, M., Zagorc, B., Brečko, J., Hiti Dvoršak, A., Moljk, B., Verbič, J., Bedrač, M., Kožar, M., Cunder, T., & Jerič, D. (2022a). Razvoj modela za sistematično spremljanje ekonomskega položaja in analizo vpliva kmetijske politike na ravni tipičnih kmetijskih gospodarstev: Zaključno poročilo. Ljubljana: Biotehniška fakulteta. Retrieved from http://www.dlib.si/details/URN:NBN:SI:doc-1J94AOK7.

Žgajnar, J., Kavčič, S., Tomšič, M., Zagorc, B., Hiti Dvoršak, A., Brečko, J., Moljk, B., & Jerič, D. (2022b). Katalog tipičnih kmetijskih gospodarstev. 1. izd. Ljubljana: Biotehniška fakulteta.

Žgajnar, J., & Zadnik Stirn, L. (2023). The use of operation research methods to support agricultural policy. In DROBNE, S. (Ed.), SOR '23: Proceedings of the 17th International Symposium on Operational Research in Slovenia. Ljubljana: Slovenian Society Informatika - Section for Operational Research, 39-42.

Appendix A – Questionnaires for the online survey

Questionnaire for stakeholder needs assessment tools

Identification

Question 1 ➔ Country/Region

If "Other" please specify the name of the Country/Region:

Question 2 ➔ Stakeholder group

If "Other" please specify the name of the Stakeholder group:

Question 3 ➔ Organisation name (optional)

Question 4 ➔ Which role best describes your relationship with participatory and multi-stakeholders tools?

If "Other" please specify the Role:

Question 5 ➔ Which participatory and multi-stakeholders tools have you been using in the design of the CSPs, and are you the most experienced with? E.g. online questionnaires, focus groups or workshops etc.

Choosing the Right Tool: Understanding Your Decision Process

Question 6 ➔ How have the following factors influenced the selection of the tool?

Question 7 ➔ Has any other factor influenced the selection of the tool?

If "Yes", please specify the name of the factor

Question 8 ➔ To what extent do you agree with the following statements (1= disagree to 5= fully agree)

Question 9 ➔ How many stakeholders have been involved in the implementation of the tool (e.g. number of participants of the workshop, focus group, consultation etc...)?

Question 10 ➔ Is the raw information provided by stakeholders traceable (i.e. it is possible to follow your data all the way back to its original source)?

If "No", please explain what the obstacles are:

Question 11 ➔ Are measures in place to avoid mistakes and redundancy in the stakeholder's responses or information provided (e.g. to verify the quality and reliability of the information)**?**

If "Yes", which ones

If "Other", please specify

Question 12 ➔ Is the tool implemented following established methodologies with known steps?

Question 13 ➔ Has the tool been used in previous programming periods of the Common Agricultural Policy (CAP)?

Question 14 ➔ Has the tool been used to assess or contribute to the following specific CAP objectives?

Question 15 ➔ Can the tool be easily tailored to assess or contribute to the assessment of the CSP to other Specific Objectives?

If "Yes", which ones?

Question 16 ➔ Which policy cycle steps has the tool been used on, and on which other policy cycle step could also be used?

Question 17 ➔ Has the tool been used to engage stakeholders from different governance levels?

If "Yes", which governance levels?

Question 18 ➔ To what extent are you satisfied with the overall ease of use of the tool? (1 = Not at all satisfied, 5 = Very satisfied)

Question 19 ➔ Does the tool/method require prior training or experience?

If "Yes", how much time you consider necessary to be able to use it properly?

Is detailed documentation concerning the underlying methodology of the tool available and accessible to all stakeholders?

If "Yes" and feasible, could you upload the documentation or provide a link?

Question 20 ➔ Do you allow feedback from stakeholders on the choice of the tools used?

Question 21 ➔ Is the feedback taken into account when the tool is used again?

Question 22 ➔ To what extent do the following factors influence the accessibility of the tool (1= not at all; 5= to a Great Extent):

Question 23 ➔ Does any other factor influence the accessibility of the specific tool?

If "Yes", please specify name of the factor

Question 24 ➔ How much time did it take for you to set up the tool? i.e time necessary to organise and prepare the implementation of the participatory tool (to identify participants, send invitations, set up the online page of the consultation or - in case of methods implemented in presence - the arrangement of the venue, etc…)

Question 25 ➔ How much time did it take to implement the process of the tool? (i.e. time necessary to implement the tool until the achievement of the intended outputs)
Question 26 ➔ How much budget is necessary for the implementation of the tool?
Question 27 ➔ In case the tool requires a subscription (i.e. monthly/annual subscription payments to use a Software), how much budget is necessary for the maintenance of the tool every year?
Question 28 ➔ How many people are necessary for the implementation of the tool?

Questionnaire for policy analysis tools

Identification
Question 1 ➔ Country/Region
If "Other" please specify the name of the Country/Region:
Question 2 ➔ Stakeholder group
If "Other" please specify the name of the Stakeholder group:
Question 3 ➔ Organisation name (optional)
Question 4 ➔ Which role best describes your relationship with policy analysis tools for ex-ante evidence-based decisions ?
Question 5 ➔ Please select the category of policy analysis tools for ex-ante evidence-based decisions you have used or you are knowledgeable about:
If "Other" please specify the name of the category of the tool:
Question 6 ➔ Please indicate the "name"of a single tool you are most knowledgeable about and the category it belongs (see previous question "Question 5" for the category).
Choosing the Right Tool: Understanding Your Decision Process
Question 7 ➔ How have the following factors influenced the selection of the tool?
Question 8 ➔ Has any other factor influenced the selection of the tool?
If "Yes", please specify the name of the factor:
Question 9 ➔ Has the tool been used in previous programming periods?
Question 10 ➔ Has the tools been used in different policy areas other than the CAP?
If "Yes", which Areas/Policies?
Question 11 ➔ Would you like to change any features or characteristics of the selected tool?
If "Yes", what would you change?
Question 12 ➔ To what extent do you agree with the following statements (1= disagree to 5= fully agree)?
Question 13 ➔ Is the input data used by the tool traceable (i.e. it is possible to follow your data all the way back to its original source)?
If "No", please explain what the obstacles are:
Question 14 ➔ Are measures in place to avoid mistakes and redundancy in the input data (e.g. to verify the quality and reliability of the information)?
If "Yes" which ones (multiple-choice are possible)?
If "Other", please specify:
Question 15 ➔ Are the processing rules for data clear and transparent? (Data management process rules refer to a set of guidelines, procedures, and standards that dictate how data should be handled, organised, stored, processed, and maintained throughout its lifecycle).
If "No", please specify the reasons:
Question 16 ➔ Has the input data sample used by the tool been demonstrated to be representative of the target population?
Question 17 ➔ Does the observation level of the data align with the desired result level?
If "No", please specify the reasons:
Question 18 ➔ Is the tool implemented following established methodologies with known steps?
Question 19 ➔ Have the results of the tool been validated?
If "Yes", which validation method? (multiple answers are possible)
If "Other", please specify:
Question 20 ➔ Is the tool regularly updated to incorporate the latest available information or feedback from end users?
Question 21 ➔ Has the tool been used to assess or contribute to the following specific CAP objectives?
Question 22 ➔ Can the tool be easily tailored in order to assess or contribute to the assessment of the CSP to other specific CAP objectives?
If "Yes", which ones:

Question 23 ➔ Which policy cycle steps has the tool been used, and on which other policy cycle step could also be used? (multiple answers allowed)
Question 24 ➔ What is the spatial coverage of the results? (i.e of the results you obtained when using it in the context of the CAP Strategic Plans) (multiple answers)
Question 25 ➔ What is the temporal extent of the results? (i.e of the results you obtained when using it in the context of the CAP Strategic Plans) (multiple answers)
Question 26 ➔ Could you please describe what kind of results are generated by the tool?
Question 27 ➔ To what extent are you satisfied with the overall ease of use of the tool? (1 = Not satisfied, 5 = Very satisfied)
Question 28 ➔ To what extent does the tool allow for changes in its methodologies, features, parameters or settings to fit in specific contexts (i.e. how flexible is the tool to changes)? (1 = Low level of flexibility, 5 = High level of flexibility)
Question 29 ➔ What is the spatial resolution of the tool? (mutliple answers)
Question 30 ➔ What is the temporal resolution of the tool? (multiple answers)
If Other please specify the name of the level
Question 31 ➔ Does the tool require prior training or experience?
If "Yes", how much time do you consider necessary to be able to use it properly?
Question 32 ➔ Is detailed documentation concerning the tool's underlying methodology, models, algorithms, and analytical techniques available and accessible to all users?
If "Yes", and feasible, could you upload the documentation or provide a link?
Question 33 ➔ Is the tool available for end-users in all Member States?
Question 34 ➔ How much time did it take for you to set up the tool? (i.e time needed for the completion of all essential tasks to make the tool operational e.g. acquisition of a software, collection of relevant data and information)
If "Other", please specify:
Question 35 ➔ How much time did it take to implement the process of the tool? (i.e. time necessary to implement the tool until the achievement of the intended outputs)
Question 36 ➔ How much budget is necessary for the implementation of the tool?
Question 37 ➔ In case the tool requires a subscription (i.e. monthly/annual subscription payments to use a Software), how much budget is necessary for the maintenance of the tool every year?
Question 38 ➔ How many people are necessary for the implementation of the tool?
Question 39 ➔ Thank you for taking the time to complete this survey. Do you have any questions or comments?

Questionnaire for policy choices supporting tools

Identification
Question 1 ➔ Country/Region
If "Other" please specify the name of the Country/Region:
Question 2 ➔ Stakeholder group
If "Other" please specify the name of the Stakeholder group:
Question 3 ➔ Organisation name (optional)
Question 4 ➔ Which role best describes your relationship with the tools adopted for the design and monitoring of the CAP Strategic Plans (CSP)?
Question 5 ➔ Which policy choices supporting tool have you been using in the design of the CSPs, and/or are you the most experienced with?
If "Other" please specify the name of the tool:
Choosing the Right Tool: Understanding Your Decision Process
Question 6 ➔ How have the following factors influenced the selection of the tool?
Question 7 ➔ Has any other factor influenced the selection of the tool?
If "Yes", please specify the factor
Question 8 ➔ Has the tool been used in previous programming periods?
Question 9 ➔ Has the tools been used in different policy areas other than the CAP?
If "Yes", which Areas/Policies?
Question 10 ➔ To what extent do you agree with the following statements (1= disagree to 5= fully agree)
Question 11 ➔ Are measures in place to avoid mistakes and redundancy in the information source (e.g. to verify the quality and reliability of the information)?
Question 12 ➔ Are the processing rules for data clear and transparent? (e.g. set of guidelines to extract information, standards that dictate how data should be handled, organized, stored, processed)
If No, please specify the reasons
Question 13 ➔ Have the results of the tool been validated?

If “Yes” which validation method?
If “Other”, please specify:
Question 14 ➔ Is the tool implemented following established methodologies with known steps?
Question 15 ➔ Has the tool been used to assess or contribute to the following specific objectives? (multiple choice)
Question 16 ➔ Can the tool be easily tailored to assess or contribute to the assessment of the CSP to other Specific Objectives?
If “Yes”, which ones:
Question 17 ➔ Which policy cycle steps has the tool been used, and on which other policy cycle step could also be used? (multiple answers allowed)
Question 18 ➔ Has the tool been used to capture different governance level? (e.g. EU level, country level, local)
If “Yes”, which levels (multiple answers are possible)?
Question 19 ➔ Could you please describe what kind of results have been generated by the tool?
Question 20 ➔ To what extent are you satisfied with the overall ease of use of the tool? (1= Not at all satisfied; 5 = Very satisfied)
Question 21 ➔ Does the tool require prior training or experience?
If “Yes”, how much time you consider necessary to be able to use it properly?
Question 22➔ Is detailed documentation concerning the underlying methodology of the tool available and accessible to all stakeholders?
If “Yes” and feasible, could you upload the documentation or provide a link?
Question 23 ➔ Is the tool available for end-users in all Member States?
Question 24 ➔ How much time did it take for you to set up the tool?
Question 25 ➔ How much time did it take to implement the process of the tool? **Question 26** ➔ To what extent are you satisfied with the overall ease of use of the tool? (1 = Not satisfied, 5 = Very satisfied)
Question 26 ➔ How much budget is necessary for the implementation of the tool?
Question 27 ➔ In case the tool requires a subscription (i.e. monthly/annual subscription payments to use a Software or to access certain data), how much budget is necessary for the maintenance of the tool every year?
Question 28 ➔ How many people are necessary for the implementation of the tool?
Question 29 ➔ Thank you for taking the time to complete this survey. Do you have any questions or comments?

Questionnaire for monitoring and data collection tools

Identification
Question 1 ➔ Country/Region
If “Other” please specify the name of the Country/Region:
Question 2 ➔ Stakeholder group
If “Other” please specify the name of the Stakeholder group:
Question 3 ➔ Organisation name (optional)
Question 4 ➔ Which role best describes your relationship with monitoring tool? (single answer)
If “Other” please specify the name of the role:
Question 5 ➔ Please select the tool categories you are knowledgeable about:
If “Other” please specify the name of the tool:
Question 6 ➔ Please indicate the "name" of a single tool you are the most knowledgeable about and the category it belongs to (see previous question)
Choosing the Right Tool: Understanding Your Decision Process
Question 7 ➔ How have the following factors influenced the selection of the tool?
Question 8 ➔ Has any other factor influenced the decision of selection of the specific tool?
If “Yes”, please specify name of the factor
Question 9 ➔ Has the tool been used in previous programming periods?
Question 10 ➔ Has the tool been used in different policy areas other than the Common Agricultural Policy (CAP)?
If “Yes”, which Areas/Policies?
Question 11 ➔ To what extent do you agree with the following statements (1= disagree to 5= fully agree)
Question 12 ➔ Is the input data used by the tool traceable (i.e. it is possible to follow your data all the way back to its original source)?
If “No”, please explain what the obstacles are:
Question 13 ➔ Is the input data kept in its original form without any alterations or modifications?
Question 14 ➔ Is the data collected from actual observations rather than being estimated or synthesized?
Question 15 ➔ Are measures in place to avoid mistakes and redundancy in the input data (e.g. data validation, double-entry verification)?

If “Yes” which ones (multiple-choice are possible)?
If “Other”, please specify:
Question 16 ➔ Are the processing rules for data clear and transparent? (e.g. set of guidelines, procedures, and standards that dictate how data should be handled, organized, stored, processed, and maintained throughout its lifecycle)
If “No”, please specify the reasons:
Question 17 ➔ Do the processing rules apply consistently to all observations?
Question 18 ➔ Is the tool following an established methodology with known steps?
Question 19 ➔ Have the results of the tool been validated?If “Yes”, how much time you consider necessary to be able to use it properly?
If “Yes” which validation method?
If "Other" please specify the validation method:
Question 20➔ Is the tool regularly updated to incorporate the latest available information or feedback from end users?
Question 21 ➔ Has the tool been used to assess or contribute to the following specific CAP objectives?
Question 22 ➔ Can the tool be easily tailored (i.e integration of new data) in order to assess or contribute to the assessment of other Specific CAP Objectives)?
If “Yes”, which ones:
Question 23 ➔ Which policy cycle steps has the tool been used, and on which other policy cycle step could also be used? (multiple answers allowed)
Question 24 ➔ What is the spatial coverage of the results? (i.e of the results you obtained when using it in the context of the CAP Strategic Plans) (multiple answers)
Question 25 ➔ What is the temporal extent of the results? (i.e of the results you obtained when using it in the context of the CAP Strategic Plans) (multiple answers)
Question 26 ➔ Could you please describe what kind of results are generated by the tool?
Question 27 ➔ To what extent are you satisfied with the overall ease of use of the tool/method? (Score 1 = Not at all, Score 5 = To a Great extent)
Question 28 ➔ What is the spatial resolution of the tool? (multiple answers)
Question 29 ➔ What is the temporal resolution of the tool? (multiple answer)
If “Other” please specify the name of the level
Question 30➔ Does the tool require prior training or experience?
If “Yes”, how much time you consider necessary to be able to use it properly?
Question 31 ➔ Is detailed documentation concerning the underlying methodology of the tool available and accessible to all end-users?
If “Yes” and feasible, could you upload the documentation or provide a link?
Question 32 ➔ Is the tool available for end-users in all Member States?
Question 33 ➔ How much time did it take for you to set up the tool? i.e time needed for the completion of all essential tasks to make the tool operational e.g. acquisition of a software, identification of data sources, organisation of the activities)
Question 34 ➔ How much time did it take to implement the process of the tool? (i.e. time necessary to implement the tool until the achievement of the intended outputs)
Question 35 ➔ How much budget is necessary for the implementation of the tool?
Question 36 ➔ In case the tool requires a subscription (i.e. monthly/annual subscription payments to use a Software or to access certain data), how much budget is necessary for the maintenance of the tool every year?
Question 37 ➔ How many people are necessary for the implementation of the tool?
Question 38 ➔ Thank you for taking the time to complete this survey. Do you have any questions or comments?

Appendix B – Definition of the sub-criteria and evaluation questions

Criteria	Sub-criteria	Sub-criteria definition	Evaluation Questions for Stakeholder needs assessment tools	Evaluation Questions for Policy choices supporting tools	Evaluation Questions for Policy analysis tools	Evaluation Questions for Monitoring and data collection tools
Accuracy	**Data Accuracy and Validity**	This sub-criterion refers to the extent to which data input and its sources are traceable and validated.	Are measures in place to avoid mistakes and redundancy and to ensure consistency in the stakeholder's responses or information provided (e.g. to verify the quality and reliability of the information)? Are results triangulated, corroborated, or substantiated by other data sources? Are different types of stakeholders involved?	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? Is it possible to easily trace back the information used for supporting the policy choices to its original source? 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)?	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 back the information used for supporting the policy choices 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?	Are measures in place to avoid mistakes and redundancy and to ensure consistency in monitoring and/or the data collection activities (e.g. to verify the quality and reliability of the collected information)? Are collected data triangulated, corroborated, or substantiated by other data sources to check validity/accuracy? Are the collected data traceable back to its original source? To what extent is the data collected based on actual/direct observations rather than on estimations?
	Tool Representativeness	Representativeness refers to the capacity of a tool to accurately and comprehensively capture or depict a system, dataset, or phenomenon. It combines the	How many stakeholders can be involved in the implementation of this sub-category of tools (e.g. number of participants of the workshop, focus group,	Are the tools able to capture the diverse facets, variations, or elements within the system or dataset it seeks to represent?	To what extent do these tools present a high degree of sensitivity (i.e. how much the results of the tools are influenced by	Are the tools able to capture the diverse facets, variations, or elements within the system or dataset it seeks to represent? What is the potential geographical/sectorial/stakeholder

		notions of accuracy and coverage in its depiction.	consultation etc.)? Does this sub-category of tools facilitate the involvement of all relevant stakeholder groups while ensuring representativeness? Is this sub-category of tools affected by biases?	Does this sub-category of tools facilitate the involvement of all relevant stakeholder groups while ensuring geographical and sectoral representativeness?	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?	coverage of these tools? Are these tools able to gather information from different sectors, stakeholders, and countries/regions?
Reliability	**Replicability**	Replicability refers to the extent to which the tool's results are (or can be expected to be) consistent across studies or other tools aimed at answering the same question, either by replicating the same approach or using new data or methods.	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?	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?	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?	Can the data collected by these tools be replicated successfully by others or through different approaches? Do the data remain consistent with other data collection activities or tools? Does collected data remain consistent over multiple data collection rounds (i.e. the collected information is always the same if data collection is replicated)?
	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.	Is the information concerning underlying methodologies, data, and results of the tool available and accessible, and is the tool accessible for external review?	Is the information concerning underlying methodologies, data, and results of the tool available and accessible, and is the tool accessible for external review?	is the information concerning underlying methodologies, data, and results of the tool available and accessible, and is the tool accessible for external review?	Is the information concerning underlying methodologies, data, and results of the tool available and accessible, and is the tool accessible for external review?
Applicability	**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.	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	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	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	To what extent can these tools be used to address different policy objectives and steps of the policymaking process? Have these tools (i.e. their collected data) 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?

			Specific Objectives and policymaking steps?	other Specific Objectives and policymaking steps?	the assessment of the other Specific Objectives and policymaking steps?	
	Experience	Experience refers to the extent to which the tool has been used across MSs, in previous programming periods, or in other policies.	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?	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?	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?	How experienced are MSs with these 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?
	Granularity	The extent to which the tool can provide results at different level of spatial (e.g. local, regional, national, and EU level) and temporal resolution.	Can the tools support analysis and policymaking at different levels, including local, regional, national and EU level? What is the level of detail that these tools can reach?	Can the tools support analysis and policymaking at different levels, including local, regional, national and EU level? What is the level of detail that these tools can reach?	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)?	Can the tools support analysis and monitoring at different levels, including local, regional, national and EU level? What is the level of detail that these tools can reach? Do the tools provide results at different temporal resolutions (e.g. results expressed in days, months, years)?
Accessibility	**Ease of use**	Ease of use refers to the extent to which the tool is easy to learn and use.	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?	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	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	Do these tools require prior training or experience, or do they require much time to be able to use them properly? Can these tools be handled by a wide range of professionals, or do they require specific professional profiles to be implemented? Is there complex technology behind the use of these tools?

				professional profiles to be implemented?	professional profiles to be implemented?	
	Tool Availability	Availability refers to the extent to which the tool is accessible, available to and usable by end users.	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?	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?	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?	Are these tools openly available to everyone, and their methodological and technical guidelines accessible to users? What are the main factors the 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?
Efficiency	**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.	How much time does it take to set up this sub-category of tools (i.e time necessary to organise and prepare the implementation, including the identification and invitation of participants, setting up the online page of the consultation or - in case of methods implemented in presence - the arrangement of the venue, etc…)? And how much time does it take to implement the process of the tool (i.e. time necessary to implement the tool until the achievement of the intended outputs)?	How much time does it take to set up this sub-category of tools (i.e time necessary to organise and prepare the implementation)? And how much time does it take to implement the process of the tool (i.e. time necessary to implement the tool until the achievement of the intended outputs)?	How much time does it take to set up this sub-category of tools (i.e time necessary to organise and prepare the implementation)? And how much time does it take to implement the process of the tool (i.e. time necessary to implement the tool until the achievement of the intended outputs)?	How much time does it take to set up this sub-category of tools (i.e time necessary to organise and prepare the implementation: e.g. installation of the software, get authorisation and credential to use it etc…)? And how much time does it take to implement the process of the tool (i.e. time necessary to implement the tool until the achievement of the intended outputs)?
	Cost-effectiveness	Cost-effectiveness refers to the extent to which the tool is able to achieve its outcomes within a reasonable amount of budget. It includes the budget necessary for patent subscriptions,	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 budget can be necessary for	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?	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?	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 budget can be necessary for the maintenance of the tool or

		maintenance and implementation.	the maintenance of the tool or subscription for patent-based tools every year?	How much budget can be necessary for the maintenance of the tool or subscription for patent-based tools every year?	How much budget can be necessary for the maintenance of the tool or subscription for patent-based tools every year?	subscription for patent-based tools every year?
	Human Resources Efficiency	Human resources efficiency refers to the extent to which the tool is able to achieve its outcomes involving a reasonable amount of human resources.	How many people are necessary for the implementation of the tool and the analysis of results? How many different professionalities are necessary for the good implementation of the tool?	How many people are necessary for the implementation of the tool and the analysis of results? How many different professionalities are necessary for the good implementation of the tool?	How many people are necessary for the implementation of the tool and the analysis of results? How many different professionalities are necessary for the good implementation of the tool?	How many people are necessary for the implementation of the tool and the analysis of results? How many different professionalities are necessary for the good implementation of the tool?

Appendix C – Benchmarking factsheets

Synthesis table of the benchmarking factsheets

No. of the factsheet	Policy step covered	Title
1	Socio-economic context and SWOT analysis	Tools for the socio-economic context and SWOT analyses of the CAP Strategic Plans
2	Needs assessment, identification and prioritisation	Tools for needs assessment, identification and prioritisation in the CAP Strategic Plans
3	Interventions setting	Tools for intervention design in the CAP Strategic Plans
4	Targets setting and financial allocation	Tools for targets setting and financial allocation in the CAP strategic plans
5	Stakeholder consultation	Tools for stakeholder consultation to inform the preparation of the CAP Strategic Plans
6	Performance	Tools for performance monitoring and review of the CAP Strategic Plans
7	Compliance	(non-IACS) Monitoring systems for beneficiaries' compliance checks
8	Compliance	(IACS-AMS) Monitoring applications for beneficiaries' compliance checks
9	Policy analyses at Eu level	EU-level modelling tools for policy analysis supporting the design of CAP Strategic Plans

The benchmarking factsheets are presented below.

Tools for the socio-economic context and SWOT analyses of the CAP Strategic Plans

OBJECTIVES AND CHALLENGES

Member States are required by Reg. 2115/2021 (Art. 115) to underpin the preparation of their CAP strategic plans with SWOT analyses of the socio-economic context for each specific CAP objective based on **common context indicators and other quantitative and qualitative up-to-date information**. For this task, most Member States have combined the (statistical) analysis of quantitative indicators with the consultation of stakeholders. The task can be challenging in that it needs to **reflect the heterogeneity across the entire national territory** (especially when there are relevant regional differences), include lessons learned from the past programming period, and properly use stakeholders' perspectives to complement the analysis of quantitative indicators.

MAIN TOOLS ADOPTED: STRENGTHS AND WEAKENESSES

World Café (DE, FI)

The world café method is a structured conversational process for **knowledge sharing** in which small groups of people discuss a topic at several small tables (like in a café). A world café can be made up of several thematic workshops, shedding light on different topics. The workshops start with short presentations related to the topic, which are followed by group discussions and idea collection (e.g. on whiteboard papers). Each group has a host that poses the questions/issues to be answered/debated and notes down the answers. At the end of each round, the participants (except for the hosts) are reshuffled into new groups, starting a new round of conversation. The host shares the insights from the previous groups with the new one. It ends with all hosts sharing the main takeaways for a joint discussion. This tool is usually used to **identify strengths, weaknesses, and development needs** related to different topics in a participatory and engaging way.

✓ It allows the involvement of a broad range of diverse stakeholders, help cross-pollinate ideas and build upon each other's contributions, and explore complex topics from multiple perspectives in a structured way.

✗ It requires the involvement of a large group of participants and intense facilitation to feed the discussion, whereas the output can tend to over-summarise the collected ideas.

Consultation forums (SE)

Consultative forums are a way to work with stakeholders, aiming to collect a wide range of views on various aspects of the strategic plans. They are used to acquire knowledge and perspectives from stakeholders on specific matters, to **increase the quality of decisions** made by strengthening dialogue, making use of expertise, collecting a broader range of perspectives, and increasing the number of involved stakeholders. The forums cover topics such as **general strategy, SWOT analysis, needs assessment, and draft interventions**. Most forums are conducted online, with participants given options to comment by speaking, writing, or sending remarks within a given time frame. This tool aims to **collect a wide range of views** on various aspects of the CAP reform.

✓ It allows for reaching a broad range of stakeholders in a cost-effective way at multiple stages of the plans' design process, in order to disseminate draft outputs from the preparation of the plans and gather a feedback.

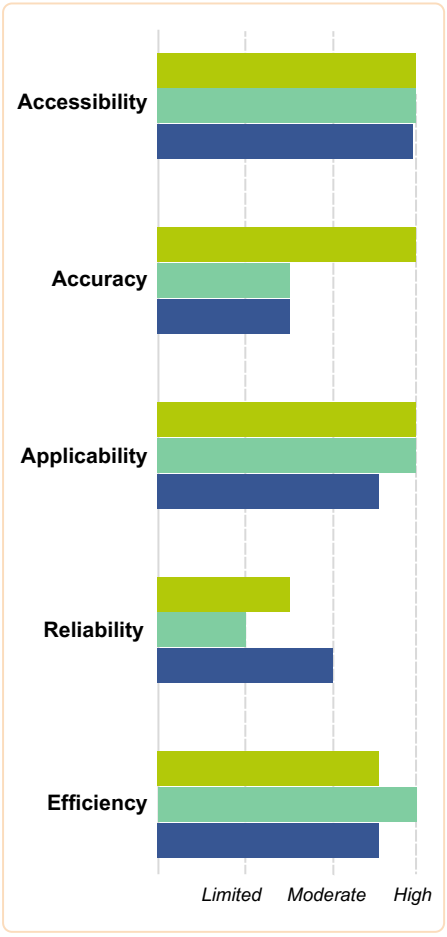
✗ Limited clarity in the selection process and lack of a co-creative approach. It is mostly a top-down, one-way feedback mechanism with limited potential to generate consensus or agreement over complex issues.

MAPP (Programmes Impact Assessment Method) (DE)

This participatory tool, based on regional workshops and qualitative analysis techniques, was used to **assess the overall contributions of the Rural Development Programmes (RDP)**. The tool, which was also recommended by European Commission guidelines and the European Evaluation Helpdesk, allows to put the program impact in relation to policy instruments and measures outside of the RDP, and to assess environmental impacts. The tool relies on a trend analysis and on the creation of an **impact matrix to assess and visualise the relations between instruments, group of instruments, and impacts**.

✓ It can look beyond the RDP by assessing other programmes. The used scoring system, while experimental, provide useful impacts' visualisations.

✗ The preparation can be time-consuming, and it was challenging to cover all instruments in one workshop. Impacts' mechanisms cannot be taken into account. It may need re-adaptation.



Find more insights in the Inventory of tools online

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.

Funded by the European Union

Tools for needs assessment, identification and prioritisation in the CAP Strategic Plans

OBJECTIVES AND CHALLENGES

Member States are required by Reg. 2115/2021 (Art. 108) to identify, assess and prioritise the needs that should be addressed within the CAP Strategic Plans. Diverse approaches and prioritisation criteria were used by Member States to conduct this task. The **hurdle of this task** relates mainly to the coherent identification of very diverse needs over the entire national territory and to the prioritisation of needs when several stakeholders are involved and might have divergent views. The task might be **more complex in regionalised countries**, where several regional authorities have their say about the needs.

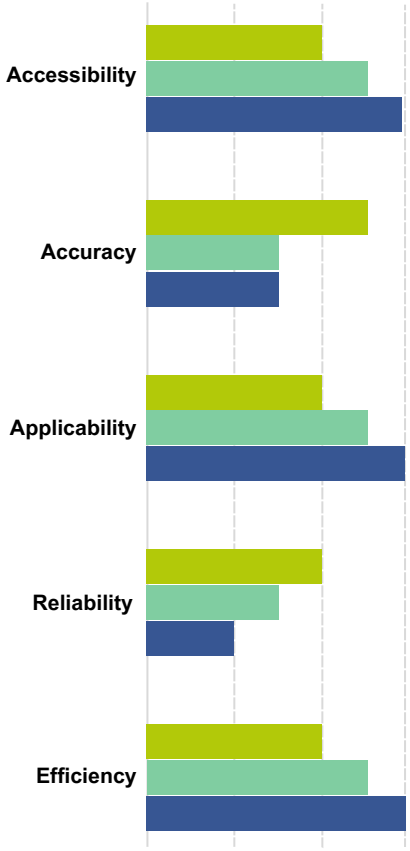
MAIN TOOLS ADOPTED: STRENGTHS AND WEAKENESSES

Multicriteria analysis (ES)

The tool provides an efficient approach **to guide policy choices and prioritise policy options**. The tool, which is based on the involvement of different stakeholders, public authorities (e.g., regional authorities), and experts, **relies on a set of pre-defined criteria and a scoring system**. The tool has been adopted particularly for the needs' prioritisation but has the potential to be used for other policy choices, such as interventions setting.

It is a relatively simple technique that enables prioritisation based on the input of a potentially large and diversified range of stakeholders, and can be easily adapted in other contexts.

It has a limited capacity to reflect the diversity of inputs in a heterogeneous territory, for example when many different regions are involved. It can be negatively impacted by subjective bias. Results from different multicriteria settings might differ.



Cumulative voting (IT, LT)

This tool consists of a voting aggregation technique to assess the prioritisation of the **identified needs and support the decision-making process** in the CSP drafting process. The tool helps formulate a shared consensus on the level of importance of each need and determine homogeneous groups of needs by the importance of intervention. It allows stakeholders to **prioritise various intervention areas** by allocating a fixed number of points across different options, thus reflecting the relative importance they assign to each area. This tool was also used to facilitate participation in needs assessment in a regionalised country. Besides, assigning weights to different groups of stakeholders can increase the level of accuracy.

It allows for the aggregation of preferences from several regions and diverse stakeholder groups, and it is easy to adopt. It ensures a participatory co-governance approach to policy.

Prioritisation is weaker when inputs are too divergent and cannot account for multiple objectives at once (trade-offs), whereas the final output might not reflect all aggregated local inputs. Participants may need prior training to understand the exercise.

Consultation forums (SE)

Consultative forums are a way to work with stakeholders, aiming to collect a wide range of views on various aspects of the strategic plans. They are used to acquire knowledge and perspectives from stakeholders on specific matters, to **increase the quality of decisions** made by strengthening dialogue,making use of expertise, collecting a broader range of perspectives, and increasing the number of involved stakeholders. The forums cover topics such as general strategy, SWOT analysis, needs assessment, and draft interventions. Most forums are conducted online, with participants given options to comment by speaking, writing, or sending comments within a given time frame. This tool aims to **collect a wide range of views on various aspects** of the CAP reform.

It allows for reaching a broad range of stakeholders in a cost-effective way at multiple stages of the plans' design process, in order to disseminate draft outputs from the preparation of the plans and gather a feedback.

Limited clarity in the selection process and lack of a co-creative approach. It is mostly a top-down, one-way feedback mechanism with limited potential to generate consensus or agreement over complex issues.

Find more insights in the Inventory of tools online

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.

Funded by the European Union

Tools for intervention design in the CAP Strategic Plans

OBJECTIVES AND CHALLENGES

One of the most important tasks for Member States is to select, design and calibrate the interventions to be included in the national strategic plans, as defined by Reg. 2115/2021 (Art. 109). This is a **controversial task** where **evidence-based decisions can conflict with political interests** and/or stakeholder views. Consequently, political negotiations usually play an important role. In some Member States, the intervention setting has benefitted from policy analyses and public stakeholder engagement, while in others the process was mostly carried out within the political backroom. The **key challenge** is to inform a political process converging to the intervention setting through scientific evidence from robust policy analyses that can strengthen the choices.

MAIN TOOLS ADOPTED: STRENGTHS AND WEAKENESSES

Eco-Scheme Farm Simulation tool (NL)

The tool consists of a farm income simulation model that offers insights into the **effect of various eco-scheme activity choices on farm income**. The tool has a twofold objective: (i) to support governments in the design of eco-schemes, and (ii) farmers in the choice of the eco-schemes to apply for. The tool has an accompanying website. **To support governments**, the tool simulates income outcomes under different eco-scheme arrangements, which inform the final design of the eco-schemes. **To support farmers**, the tool allows them to choose from a menu (22+ eco activities) and calculate a performance-related payment per hectare: farmers can simulate different choices and see what per hectare payments they generate. The tool was developed because of the particular eco scheme functioning in the Netherlands, where the eco-scheme payments are based on a point system.

It can support both policymakers in the design phase, and farmers' choice. It is openly available and easy to interpret.

It was specifically tailored to the Dutch case, based on a point system for eco-schemes: it might require relevant adaptations with other eco-scheme systems.

Farm Income calculation tool (NL)

The Farm Income Calculation Tool is based on FADN data and offers insight into the **effect of various policy options on farm income**. The tool has supported the analysis of various policy options, especially direct payments and basic income support, and their potential effect on farm income. The tool takes into account different farm sectors and different farm size classes. Thus, it is relevant mainly in the context of interventions setting.

It is based on a reliable and rich data source available in all Member States, and it is relatively easy to use use and interpret.

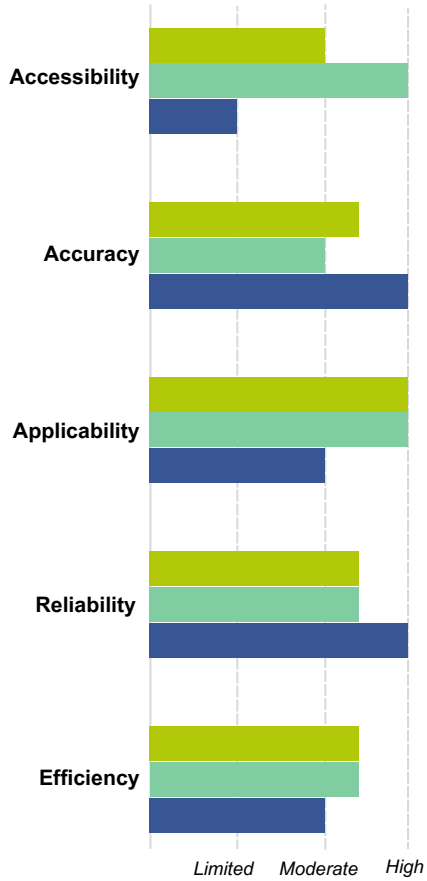
The analysis is limited to direct payment interventions, and focuses only on income outcomes, without considering environmental and social implications and trade-offs.

Eco-scheme modelling tool (DE)

The tool is based on a sophisticated approach based on three distinct models to **address the multifaceted aspects of eco-scheme design**. It conducts initial evaluations of ecological and economic impacts according to farm size and type, a **projection of the ensuing budget requirements** and, in addition, a specific estimation of the costs associated with farming without chemical-synthetic plant protection products for a year. The models were developed **as part of the ex-ante analysis activities**.

It provides a comprehesnive view on the different implications of adopting different eco-schemes.

Different models need to be calibrated to investigate different eco-schemes, whereas results and indications are effective in general scenarios but less in very specific scenarios.



Find more insights in the Inventory of tools online

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.

Funded by the European Union

Tools for targets setting and financial allocation in the CAP strategic plans

OBJECTIVES AND CHALLENGES

As by Reg. 2115/2021 (Art. 112), Member States should decide on the financial allocations for the different interventions and objectives and on the targets to be achieved. On the one hand, it is challenging to properly calibrate the allocations to different interventions, and to set targets that are at the same time ambitious and realistically achievable (as this has direct implications on Member States' accountability). On the other hand, there is usually **high pressure** from stakeholders and civil society at this stage. A crucial challenge for this task is not only to support complex technical decisions with scentific evidence, but also to **ensure the consistency between needs, interventions, targets and budget**.

MAIN TOOLS ADOPTED: STRENGTHS AND WEAKENESSES

Support Simulation Tool (BE-W)

The Support Simulation Tool acts as an aid and income simulator, designed to **evaluate the effects of financial assistance** based on the technical and economic features of farms, considering factors such as farm size and area. It evaluates direct payment support across different farm types, comparing it with 2019 data for each farm type. The tool's functioning is further enhanced by incorporating FADN data to project the effects on income. The tool provides numerical forecasts detailing the **support level according to different criteria** such as technic economic orientation, economic size, region, and farm size. It also offers aggregate projections that combine different types of farms.

It is based on FADN datasets which are available in all EU. It captures the diversity across farms, and it is easy to interpret.

This tool is tailored to the Wallonia context, and might need adaptations to other contexts. It requires input and supervision by experts, whereas the outcome is limited to farm income, without considering environmental and socila implications.

I/O/I Matrix (NL)

The tool is a primarily qualitative tool (indicating direction of impacts, sometimes with some rough quantification on impacts and budget costs) that represents on the on the one hand a **detailed list of objectives** and on the other hand a **detailed set on instruments** (covering all the relevant articles of the CAP Directive). The matrix-cells show the impact of instruments on the objectives. As such it can **support policy makers in the choice of an effective policy mix**. It has been used to carry out an analysis of possible policy options for achieving the goals set in the CSP (making use of 4 scenario's), and the integral impact of those options as regards the economic, ecological and social objectives set and their interrelations.

It allows rapid computation and it is flexible in its specifiфcation, and a large set of policy options can be studied.

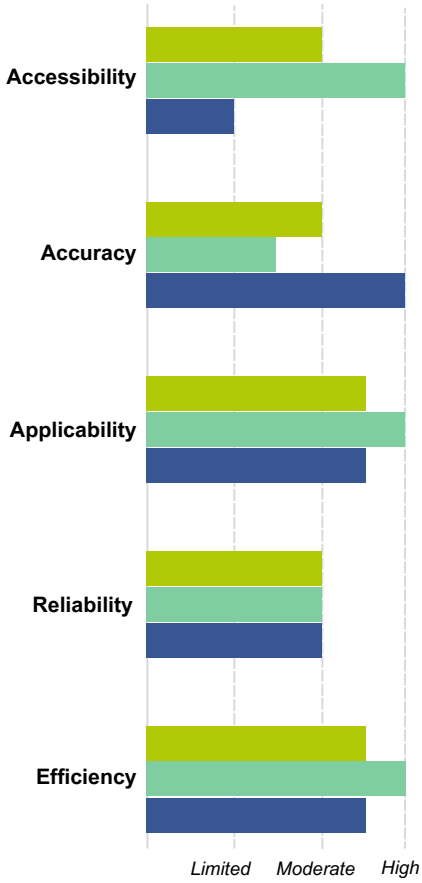
It is based on strong assumptions intended to guide policy choices and implying a simplification of reality, and is restricted to the production side of the economy.

SitFarm (Slovenian) Modelling Tool (SI)

The tool enables the **assessment of different policy scenarios** on economic and environmental performance of farms in national context, and is used for monitoring, financial planning, and agricultural outlooks. It provides nuanced **analyses across various agricultural levels**, from individual holdings to subsectors and agriculture as such, with impact assessments of direct payments, eco-schemes and area with natural constraints measures. It consists of a static farm model, with different modules taking the form of MS Excel spreadsheets. It includes models of typical agricultural holdings, model (product budget) calculations, and a farm (programming) model.

It is adaptable for use at different spatial scales, from sub-national levels with additional data to higher spatial/governance scales. The approach is easily replicable in other contexts, and it is easy to interpret.

It does not yet support modeling for agri-environmental-climate payments, investment support, and other interventions. It requires ad-hoc survey-based data sources which are not available in all Member States.



Find more insights in the Inventory of tools online

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.

Funded by the European Union

Tools for stakeholder consultation to inform the preparation of the CAP Strategic Plans

OBJECTIVES AND CHALLENGES

Member States are required by Reg. 2115/2021 (Art. 106) to involve a broad range of stakeholders and authorities in consultation activities. The main challenges for this task are to define which steps of the design process should be informed by consultation activities, and the **proper engagement approach (**open versus restricted, co-creative versus informative, bottom-up versus top-down, or a combination of them), while ensuring the transparency of the process. In fact, several engagement strategies are available to involve stakeholders, each with different implications. Generally speaking, the choice of engagement tools should be a**djusted to the policy design step** that is informed by the consultation.

MAIN TOOLS ADOPTED: STRENGTHS AND WEAKENESSES

Online consultation platforms (Impactons (FR) and Otakantaa.fi (FI))

Generally based on designated online platforms, online consultations are dissemination and opinion collection strategies aiming at disseminating to and gathering diverse perspectives from a vast audience. This tool can combine **social media dissemination** and **feedback monitoring** with online **questionnaires** and dedicated mailboxes to collect opinions. The consultations can be permanent all along the plans' preparation process and/or focus on specific steps of the plans' design.

It allows the involvement of a very large and diverse audience, through multiple complementary virtual channels, that is not reachable with traditional consultation methods. Considering the potential audience, it is relatively cost-effective.

The accessibility to digital technologies differs among stakeholders, possibly leading to an unbalanced audience composition, whereas bigger interest groups might self-select. Processing large amounts of qualitative information can be complex and time-demanding, and the inputs' origin is hardly traceable.

Town Hall Meetings (IR)

The tool consists of meetings that are conducted throughout the country on various dates, offering an **accessible platform for stakeholders to share their views, concerns, and suggestions** regarding CAP reforms. Additionally, these meetings can take place virtually. Initially, an informative presentation on the consultation topic is delivered, succeeded by one or **more Q&A sessions** about the topic, during which participants' feedback is gathered. Public consultation documents containing specific questions related to the topic are also distributed to participants. Their responses and feedback are then collected through additional means such as email or mail.

It enables the engagament of a large and diverse range of stakeholders throughout the whole territory, whereby different local perspectives can be captured.

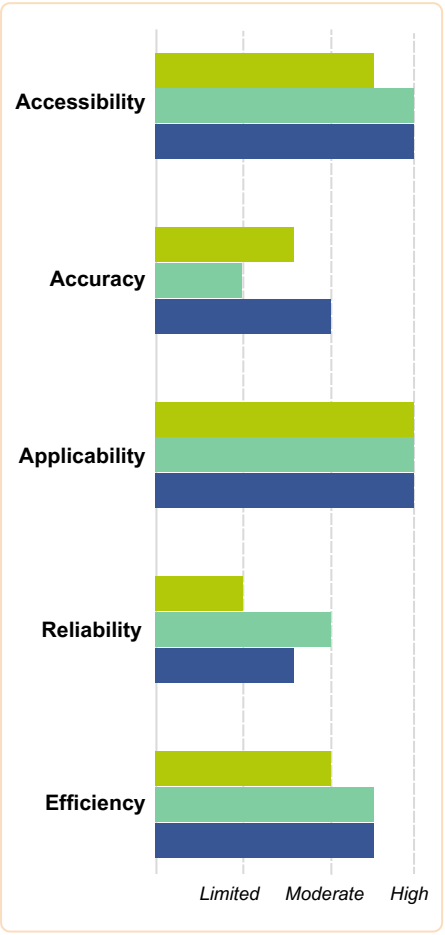
The processing of the collected data can be very time-demanding, and setting up a structured approach for a meaningful analysis can be difficult.

World Café (DE, FI)

The world café method is a structured conversational process for **knowledge sharing** in which small groups of people discuss a topic at several small tables (like in a café). A world café can be made up of several thematic workshops, shedding light on different topics. The workshops start with short presentations related to the topic, which are followed by group discussions and idea collection (e.g. on whiteboard papers). Each group has a host that poses the questions/issues to be answered/debated and notes down the answers. At the end of each round, the participants (except for the hosts) are reshuffled into new groups, starting a new round of conversation. The host shares the insights from the previous groups with the new one. It ends with all hosts sharing the main takeaways for a joint discussion. This tool is usually used to **identify strengths, weaknesses, and development needs** related to different topics in a participatory and engaging way.

The tool allows the involvement of a broad range of diverse stakeholders, help cross-pollinate ideas and build upon each other's contributions, and explore complex topics from multiple perspectives in a structured way.

It requires the involvement of a large group of participants and intense facilitation to feed the discussion, whereas the output can tend to over-summarise the collected ideas.



Find more insights in the Inventory of tools online

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.

Funded by the European Union

Tools for performance monitoring and review of the CAP Strategic Plans

OBJECTIVES AND CHALLENGES

The new CAP shifts from a compliance-based model to a performance-oriented one. Consequently, Member States are now required to monitor the performance of their strategic plans (Reg. 2115/2021, Art. 128). While monitoring systems are already in place since the previous programming period (especially for compliance checks and assessment of the second Pillar of the CAP), Member States need to **enlarge the set of data collected to monitor the progress** towards several economic, environmental and social indicators. It can be challenging for Member States to identify proper data for certain indicators (especially at farm level, and for certain social and environmental indicators), whereas **existing databases are often not interoperable** and their data are not yet or smoothly available for performance review.

MAIN TOOLS ADOPTED: STRENGTHS AND WEAKENESSES

Digital Farm Book - Irrigation Calendar (PT)

The tool Is a decision-making tool **for crop irrigation**. It collects information on irrigation systems to evaluate efficiency, analyse crop water requirements, and assess legal constraints. It is supported by two other tools that feed information on **agrometeorologic conditions** and irrigation to assess the crop's water status according to FAO's methology. Outputs include detailed information about crop water requirements, irrigation timing, legal constraints, and consumption data, all of which can be accessed online.

✓ It supports farmers' irrigation decisions, and allows farm-level data collection for legal compliance check and performance monitoring.

✗ The tool is applicable only to monitor the irrigation water efficiency measure.

Dashboard Configuration Tool (NL)

The tool consists of a configuration feature to **monitor progress on various national-level result indicators**. It is a dashboard that compiles this data and generates a progress report. The managing authority utilises the configuation tool to monitor progress with regards the various result indicators in the Dutch context, including both economic, environmental and social indicators. Based on this dashboard, managing authorities perform quantiative analysis to assess the progress towards targets and provide qualitative explanations to the results in case certain targets are not achieved. The dashboard, therefore, **supports the drafting of the annual performance reports** required by regulation.

✓ It allows monitoring of several results indicators (potentially all types of indicators) at national level, and can provide useful visualisations.

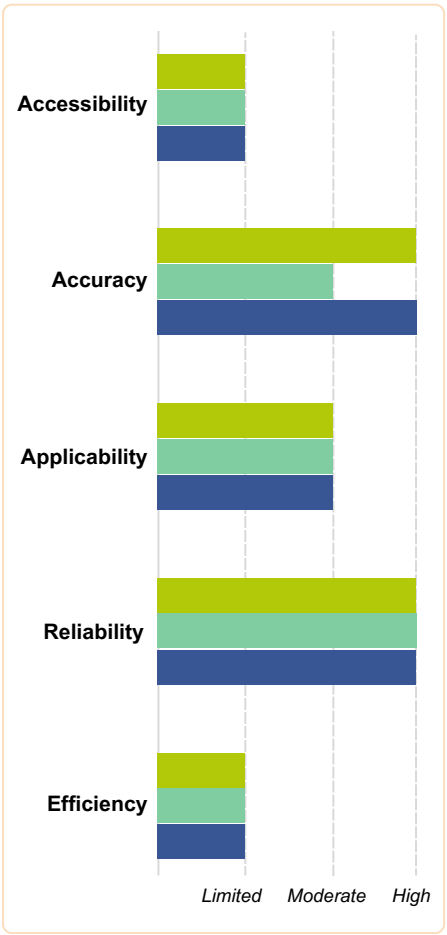
✗ It provides reports and visualisations, but analyses are not automised (i.e. they are conducted independently).

Synapse (FR)

It is a consolidated **IT architecture** capable of streamlining, auditing, and validating data for the Annual Performance Report, ensuring **uniqueness among CAP beneficiaries** across different paying agencies. Data sent by each paying agency undergoes checks for completeness, authorisation, and correctness before processing.

✓ The tool emphasises interoperability between various management systems, minimising redundancy in data requests, and establishing a unique identifier number (GUID).

✗ The accuracy of the data depends on the frequency of the reporting activities by the paying agencies.



Find more insights in the Inventory of tools online

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.

Funded by the European Union

(non-IACS) Monitoring systems for beneficiaries' compliance checks

OBJECTIVES AND CHALLENGES

As in the previous programming period, Member States are required to monitor the compliance of beneficiaries, also considering that new interventions (e.g. eco-schemes) and requirements (e.g. social conditionality) have been introduced. Yet, Member States are now called to **reduce the burden for farmers** and the administrative costs fostering the simplification of the monitoring systems, while improving their **ability to collect data at farm level**. All Member States largely rely on IACS systems to collect data about beneficiaries' compliance. Nonetheless, there are other (non-IACS) monitoring systems that can be established and complement IACS.

MAIN TOOLS ADOPTED: STRENGTHS AND WEAKENESSES

ClassyFarm (IT)

It operates as an integrated system for **categorising breeding based on risk factors**, utilising scientifically validated coefficients to convert collected data into a numerical risk indicator for farms. Methodological transparency is maintained through the public disclosure of the risk assessment and categorisation processes. One of the fundamental features is its **interoperability with other IT systems** such as the Livestock Registry, the Electronic Veterinary Prescription and thesanitary diagnostic laboratories.

- It enhances the efficacy of data by making interoperable already existing administrative data sources.
- It focuses only on specific areas, such as animal welfare, farm biosecurity, antimicrobial consumption, and antimicrobial susceptibility profiles.

LAD IT System (LV)

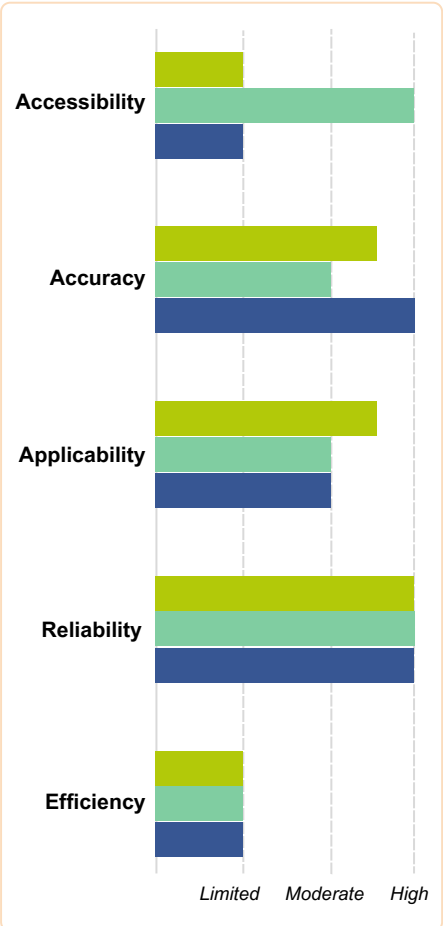
The tool is an IT system with **multiple subsystems and data exchange services** accessible to stakeholders, including the RSS, Ministry of Agriculture, and evaluators. It handles **extensive daily operations** and enables evaluators to access applicant data from other systems. Its workflow involves assessing needs, capturing and verifying data, structured management, and integrating with Oracle for efficient reporting to prevent system overload.

- Aknowledged efficiency and effectiveness in managing agricultural data.
- Technical and technological complexity.

Digital Farm Book - SIEX (ES)

This tool is designed to collect, store, and exchange information not only on the structure of the agricultural holding but also on the agricultural practices—from ploughing or harvesting to fertilizing and pesticides applications—that are performed on the parcel. The data in this tool provides farmers with a deeper understanding of their **operation's performance**. Additionally, when this information is shared with the competent authority, it can serve as **evidence of compliance** with certain CAP requirements.

- The tool provides farmer with valuable information on their holding's performance and it assists in verifying that they meet the requirements for receiving aid.
- The tool primarily contributes to environmental objectives (SO4, SO5, SO6), and the cross-cutting objective, with unclear ideas of its contribution to the remaining objectives.



Find more insights in the Inventory of tools online

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.

Funded by the European Union

(IACS-AMS) Monitoring applications for beneficiaries' compliance checks

OBJECTIVES AND CHALLENGES

As in the previous programming period, Member States are required to monitor the compliance of beneficiaries, also considering that new interventions (e.g. eco-schemes) and requirements (e.g. social conditionality) have been introduced. Yet, Member States are now called to **reduce the burden for farmers and the administrative costs** and fostering the simplification of the monitoring systems. IACS systems are in place in all Member States. However, new monitoring applications and technologies are available that could help enhance the monitoring system. The diversity of technologies adopted across the EU represents an **opportunity for peer-learning and replication** of good practices.

MAIN TOOLS ADOPTED: STRENGTHS AND WEAKENESSES

Agrisnap (EL)

The tool is a **precision photo app** that allows farmers to send pictures and information about their land parcels to a grant payment body for validation of payments. Through the mobile application, farmers can take geotagged photos of the parcels, which **helps prove the land uses and calculate the aids that will be paid**. The key features offered by the app make the process easier for users.

- It facilitates exchanges between agecies and farmers or advisors.
- Frequency of updates or improvements could affect accuracy of the tool.

iSIP -Parcel Identification (PT)

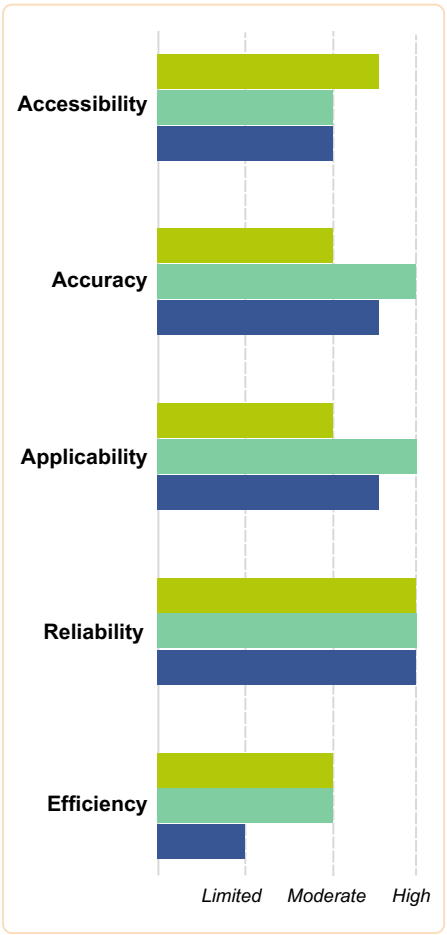
The tool is a public access online platform and a technology-based tool. It comprises a database of georeferenced datasets used for **annual checks on agricultural parcels and farming practices**, updating the database, and allowing farmers to confirm or refute collected data. It includes information on **legal constraints, agri-environmental measures, eco-schemes**, and relevant information for rural property management.

- It allows updated field information collection, which can be confirmed, rejected and exported by farmers through the app.
- The accuracy of data collected depends on the quality of data sources and on the validation by farmers.

BiedjaCam (MT)

The tool streamlines agricultural processes, provides essential information, and **acts as a digital hub for farmers**. It enables instant communication between farmers and agencies, allowing for the submission of geotagged images for evidence-based verification, thereby **reducing the need for physical on-site checks**.

- it allows farmers to rectify issues and claim support scheme payments in a timely manner.
- It is used only in the monitoring of particular AECCs and Eco-schemes.



Find more insights in the Inventory of tools online

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.

Funded by the European Union

EU-level modelling tools for policy analysis supporting the design of CAP Strategic Plans

OBJECTIVES AND CHALLENGES

There is long experience in conducting analyses of the CAP at the EU level. Several different models have been set up by the Joint Research Centre or by Universities and research institutes to analyse different aspects of EU agriculture, which can **cover one or multiple countries**, and the global agricultural and trade system. These models, which can be used in an integrated way, can provide valuable evidence to **assess different policy impacts**. Besides the Joint Research Centre, a broad network of experts is available in Europe to set up, maintain and implement these models. While technically complex to set up in-house, Member States have the possibility to complement their own analyses with those provided by these experts.

MAIN TOOLS ADOPTED: STRENGTHS AND WEAKENESSES

CAPRI

The tool is a global agro-economic modelling system designed for **assessing economic and environmental impacts** on agriculture at the regional level. It relies on data from EUROSTAT, FAOSTAT and FADN, offering projections for medium and long-run perspectives. CAPRI has been widely implemented for the **analysis of the CAP**.

✓ It can allow to establish linkages on environmental issues within the CAP and other trade policies.

✗ It cannot completely account for the micro-level heterogeneities within the analysed regions. Its use requires specialised expertise.

AGMEMOD

The tool is a multi-country, multi-market model that covers all Members States, some non-EU countries and a stylised version of the rest of the world. The model provides **annual projections for markets of the main agricultural commodities** at national and aggregated EU levels and analyses the impacts of agricultural policies. It relies on EUROSTAT, FAOSTAT, ITC statistics, and other national statistics.

✓ It can assess agricultural policy impacts in each Member State and illustrate the differing impacts of policy changes across the EU.

✗ It may not fully capture long-term policy or market shifts. Calibration is complex and requires experts.

IFM-CAP

The tool is a **static optimisation individual farm-level model assessing CAP economic and environmental impacts on farms**. It simulates farms' behaviour based on FADN data and can be applied to all EU. It originates from the increasing demand for a micro-simulation tool modelling farm-level impacts and capturing the heterogeneity across EU farms.

✓ It captures farm heterogeneity and relies on FADN data which are available in all EU-27. It considers both economic and environmental outcomes.

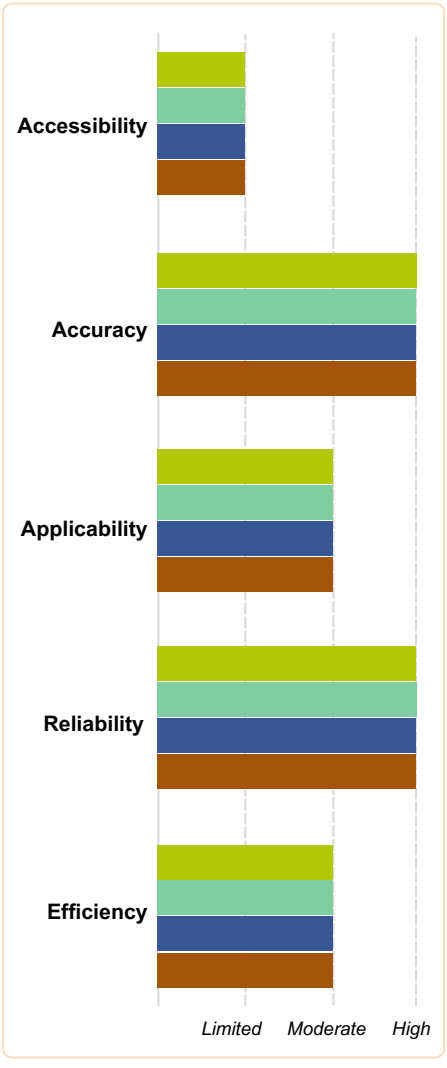
✗ It relies on strong assumptions, whereas farm exit/entry decisions are not modeled, and interactions between agricultural holdings are ignored.

GLOBIOM

The tool is a global economic model for both agriculture and forestry. It relies on major statistical databases (FAOSTAT, EUROSTAT, FADN) and **high spatial resolution bio-physical databases**, which make it particularly suitable for assessment of economic, environmental and climate impacts of policy (including carbon emissions and sinks, biodiversity, and water). It has been widely used in **policy impact assessments** by the European Commission, and other non-EU countries.

✓ It ensures a consistent and detailed coverage of both agricultural and forestry sectors. A holistic approach to land use-related questions.

✗ EU agriculture representation is undergoing a major overhaul following agricultural policies reforms.



Find more insights in the Inventory of tools online

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.

Funded by the European Union