D3.1 Non-linear dynamic model of the farm agents

agricore

Deliverable Number	D3.1
Lead Beneficiary	IDE
Authors	IDE, UNIPR, AKD
Work package	WP3
Delivery Date	September 2022 (M37)
Dissemination Level	Public

www.agricore-project.eu

Document Information

Project title	Agent-based support tool for the development of agriculture policies
Project acronym	AGRICORE
Project call	H2020-RUR-04-2018-2019
Grant number	816078
Project duration	1.09.2019-31.8.2023 (48 months)
Deliverable authors	IDENER Team
Deliverable reviewers	Lisa Baldi (UNIPR), Waldemar Bojar (PBS), Piotr Baranowski (IAPAS), Peyman Uysal (AKD)

Version History

Version	Description	Organisation	Date
0.1	Deliverable template proposal	IDE	20 Jul 2022
0.2	Modifications to ToC and template approval	UNIPR	31 Jul 2022
0.3	First (incomplete) draft	IDE	10 Sep 2022
0.4	First round of revision and comments	UNIPR	17 Sep 2022
0.5	Second (complete) draft ready to be checked	IDE	15 Oct 2022
0.6	Revisions and comments	UNIPR, IAPAS, AKD, PBS	22 Oct 2022
0.7	Implementation of changes and corrections	IDE	24 Oct 2022
0.8	Final exportation and formatting procedures	IDE	24 Oct 2022
1.0	Final Version complete	IDE	25 Oct 2022

Disclaimer

All the contributors to this deliverable declare that they:

- Are aware that plagiarism and/or literal utilisation (copy) of materials and texts from other Projects, works and deliverables must be avoided and may be subject to disciplinary actions against the related partners and/or the Project consortium by the EU.
- Confirm that all their individual contributions to this deliverable are genuine and their own work or the work of their teams working in the Project, except where is explicitly indicated otherwise.
- Have followed the required conventions in referencing the thoughts, ideas and texts made outside the Project.

Executive Summary

AGRICORE is a research project funded by the European Commission under the RUR-04-2018 call, part of the H2020 programme, which proposes an innovative way to apply agent-based modelling to improve the capacity of policymakers to evaluate the impact of agricultural-related measurements under and outside the framework of the Common Agricultural Policy (CAP).

This deliverable presents the general description of the AGRICORE agent-based model and the AGRICORE agent-based simulation.

In the introductory section, the purpose and usefulness of building this model is justified, and the rudiments of agent-based modelling and the advantages it can offer for analysing the impact of public agricultural policies are briefly presented.

This is followed by a description of the model using the ODD+D protocol, which facilitates its understanding by modellers and users and allows for easy comparison with other existing models. The fourth section presents the task flow associated with the execution of an impact analysis using the AGRICORE tool. It includes descriptions of the ABM phases (Initialisation, Calibration and Simulation itself), including their different sub-stages, and mentions the two additional phases (Calculation of KPIs and Visualisation of results).

The fifth section presents the equations of the models (own or external) that allow updating the state of the agents after each simulation step (at the end of each of the simulated agricultural campaigns) according to the actions they have implemented and according to the perturbations external to them, which logically they cannot control. These models can also be used for decision-making, as explained in deliverable D3.2 of this same AGRICORE project.

Finally, conclusions are stated in section 5.

Abbreviations

Abbreviation	Full name
ABM	Agent-based model
ABMS	Agent-based model and simulation
ABS	Agent-based simulation
AES	Agri-environment scheme
AH	Agricultural Holding
AI	Artificial Intelligence
EC	European Commission
FADN	Farm Accountancy Data Network
IAM	Impact Assessment Module
ICT	Information and Communication Technologies
LT	Long-Term
MERRA	Modern-Era Retrospective analysis for Research and Applications
MPC	Model Predictive Control(ller)
MS	Member States
ODD+D	Overview, Design and Details + human decision-making
PMP	Positive Mathematical Programming
SPEI	Standardised Precipitation and Evapotranspiration Index
SPG	Synthetic Population Generation (Generator)
ST	Short-Term
WD_{min}	Minimum withdrawal for family support

List of Figures

List of Tables

Table of Contents

1 Introduction

The global objective of the European Commission linked to the Rural Renaissance call (H2020-RUR-2018-2020) within which AGRICORE was granted is to develop and maintain appropriate instruments for use in the design of policies influencing the agricultural sector and rural areas, as society assigns an increasing number of objectives, in addition to securing food production, to these policies (H2020 Call Description). Relying on models and tools able to inform the design of these policies and to monitor their effects is part of the UE strategy for evidence-based policy-making [1]. The call directly ask proposals to take advantage of new socio-economic approaches and increased possibilities opened up by progress in the information and communication technologies (ICT) area, specifically mentioning that agent-based approaches might be advisable to ensure the focus on local effects of global events and EU policies. That is exactly the aim of the AGRICORE suite of tools, including the AGRICORE agent-based model.

Agent-based models (ABMs) are models where individuals (so-called agents) are described as unique and autonomous entities that usually interact with each other and their local environment [2]. An agent may represent not only a physical (animal or human) individual, but any kind of generic entity (a company, an institution, or any other object) that pursues certain goals.

The typical elements of an ABM are listed below [3] and graphically represented in Figure 1.

- A set of **agents**, each one of them defined by its attributes and methods. The attributes of an agent are the set of properties that define it and differentiate it from other agents. They generally reproduce observable and distinguishable characteristics of the real element represented by the agent.

  The attributes of an agent take on value throughout its simulated life, and this value can be static (constant throughout the simulation) or dynamic (variable throughout the simulation). The value of some attributes can be shared (identical) for all agents locally or globally, but at the conceptual and software implementation level, each agent stores its own register with the value of its attribute. The set of values that the attributes of an agent take at a given time instant defines the overall internal state of that agent at that time instant. Agents of the same type share the same attribute structure, differing in the value taken by the attributes for each agent; agents of different types differ both in the attribute structure (although they may share a set of attributes) and in the values taken by the attributes. The methods of an agent are the algorithmic routines that determine and implement its potential actions. They include those routines that determine agent's behaviour and those which link the agent's situation with its action or set of potential actions (e.g. the method that an agent uses to identify its neighbours). The behaviour of the agent is computed inside the corresponding method using a set of rules which describe how states are translated to actions or new states. These rules can also be static or dynamic (in case agents are prone to learning). The observable result of an agent's behaviour are its actions, the actual activities that it performs based on the application of decision rules on its states. Non-action should be considered a behavioural outcome as well. Actions can be directed at other agents but also can be triggered as a response to external stimuli, either from other agents or from the environment.

- A set of **relationships,** given as a topology of connections defining how and with whom the agents interact. An agent can relate directly to other agents, of the same or different types. It can also relate indirectly to other agents through aggregating elements that are represented in the ABM as independent modules (e.g. markets in which the individual actions of all agents have an impact on the determination of an overall price for all). Finally, each agent also relates to the environment it inhabits.

- The **environment** of the agents is the physical or virtual context in which the agents "live". Agents cohabiting with an agent in the same environment can also be considered as part of its environment, understood in a broad sense. The environment contains all the information external to the agent used in the decision-making processes and provides a structure or space for agent interaction. Agents can affect the environment and be affected by it as a consequence of the specific rules they use to determine their actions [4].

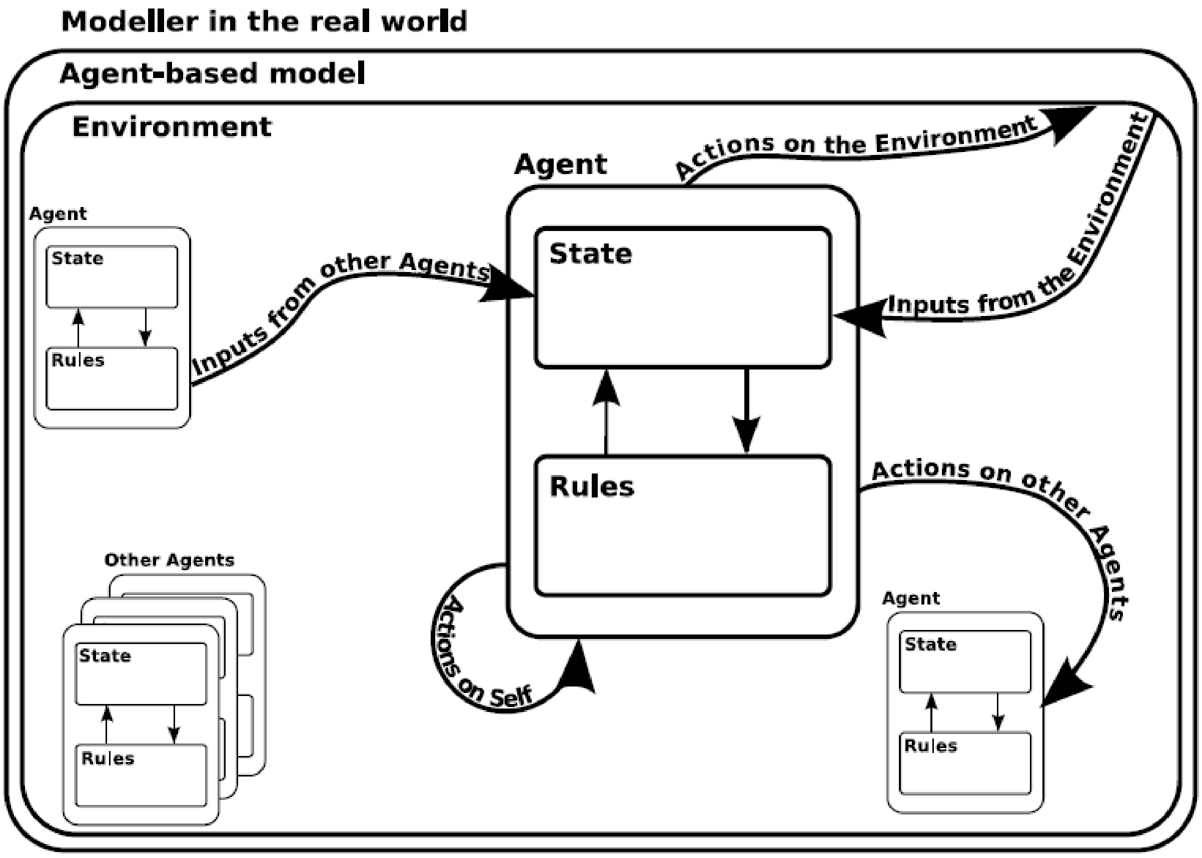

Figure 1: Structure of an ABM. Taken from [4]. Actions on self are those actions initiated by the agent that change its own state. Actions on the environment represent the effect of agent's activity on the common environment shared by all agents. Not all agents interact between them, that's why there is a subset of 'Other agents' with no interacting arrows.

The essential characteristics of agents in ABM models are the following [2][5]:

- Agents are self-contained (their actions are fully determined by the application of their rules on their states without the need for external directions), modular (have a clear boundary that defines whether something is part of an agent, is not part of an agent, or is a shared attribute), and uniquely identifiable (they can be distinguished from and recognized by other agents).
- An agent is autonomous (can function independently of its environment and of interactions with other agents at least over a limited range of situations) and self-directed (pursue its own objectives).
- An agent is social (the behaviour of other agents, sensed through interactions, influences the self-behaviour of the agent). These interactions may occur locally (with a vicinity of agents determined by spatial distance, connection network or other criteria) or globally (with all the agents of the whole simulated environment).

- Agents are active (they can initiate actions to achieve their internal goals, rather than merely waiting passively for other agents or the environment to govern their acts).
- An agent has a time-dependent state. Time can be considered as a part of the environment. While real elements act in parallel in continuous time, an agent-based model is forced to be simulated in the discrete-time of computers. This is reflected by the use of a tick as the smallest unit of time between which the dynamic changes of the agent's states can be resolved. Simulations can play with discrete time by redefining how much time a tick is meant to represent.

Using ABMs allows for addressing problems that concern emergence, i.e. system dynamics that arise from how the system's individual components interact with and respond to each other and their environment. Hence, ABMs delve into how a system's behaviour arises from (and is linked to) the characteristics and behaviours of its individual components. ABMs are useful for understanding emergence problems because they have an across-level approach. In particular, they are concerned with (at least) two levels and their interactions: these models are interested in what happens to the system according *to what the individuals do and what happens to the individuals because of what the system does.*

Although it is not the objective of this deliverable to make an exhaustive analysis of its applications, agent-based modelling and simulation (ABMS) has been used in a wide range of cases spanning the physical [6], biological [7], social [8], and management sciences [9]. In the field of policy design and evaluation, ABMS have also been used for the analysis of tobacco regulations [10], water domain policies [11], economic policy [12], and energy policies [13].

ABMS also has a remarkable track record in terms of its application to agricultural policy analysis. Kremmydas [14] does an excellent review of papers claiming to use ABM for agricultural policy analysis. In an earlier working paper [Kremmydas2012], he also makes a comparison of the most popular ABMS up to that date: Agripolis, RegMAS, MP-MAS and SWISS-Land. In this same working paper he compiles some of the advantages that various authors have argued that ABMs show over traditional modelling methods for agricultural policy analysis:

- Bottom-up vs. top-down approach: traditional modelling mechanisms compute the effect of agricultural policies at the sectoral level and at a high geo-spatial level (e.g. country level) and then have to design methods to 'disaggregate' this overall effect by regions or counties and by farm types (top-down approach). These methods are often not fully grounded at the theoretical level and are subject to errors. Meanwhile, ABMS follow a bottom-up approach in which the individual behaviour of agents and their interactions are simulated to derive the macroscopic equilibrium, which can bring better results.
- Ability to catch "emergent phenomena": ABMS can adequately mimic the non-lineal or even chaotic behaviour of a complex system such as the agricultural sector, formed by thousands of individual and heterogeneous agricultural holdings. The classical algebraic or analytical methods cannot easily include the complexity of those systems and have difficulties computing their final state. ABMS are also useful even when the system exist far from any type of equilibrium [15].
- Heterogeneity and rationality: ABMS allows for much easier modelling of heterogeneous and not fully rational entities [16]. The agricultural sector, subject to external public policies, is a complex adaptive system ideal to be analysed through ABMS as it shows a high degree of heterogeneity, not only in terms of the typology of the entities that form it (e.g. agricultural holdings of different economic size or different type of farming may differ in their attribute structure or in the range of values they can take) but also in terms of the rules that represent the way in which they decide their actions (e.g. the behaviour of familiar farms may be driven by different rules than the ones of commercial farms). In agent-based models, bounded rationality is usually assumed as a consequence of agents facing limited information and/or

information processing capacity and finite resources. Additionally, adaptative mechanisms and learning capabilities are features that agents can be endowed with.

- Easiness to model time and spatial sparsity: spatial diversity can be considered part of the general heterogeneity of agents, but ABMS allows to incorporate the expected spatial autocorrelation of the attributes of agents (agricultural holdings) located in the same geographical vicinity and to investigate the spatial dynamics of various properties in a simpler way. The spatial placement of farms can be included in the population synthesis process during the initialisation of the ABMS. The time dimension is usually ignored or downgraded in conventional modelling approaches. The macro-behaviours that emerge from interactions between agents can be fed back to influence their future actions, but this requires contemplating the different temporality of these behaviours. Also, the policy-makers might be interested not only in the effects of certain policy once it reaches its stationary regime but also during the transient period between its entry into force and the time when the majority of its potential beneficiaries have effectively acceded to it and perceive its effects.

However, ABMS also show a set of drawbacks. Axtell [15] states that the robustness of ABMS is weaker than the one of conventional mathematical methods, as its solutions are strongly dependent on the initial conditions of the simulation. The computation of sensitivity analyses within ABMS is also trickier and more complex. Another issue is referred to as the "black box" criticism: since ABMS is implemented in a programming language, it is challenging to explain the model internals, such as the assumptions and the specific algorithms behind it, in a consistent and understandable manner. The use of protocolised methodologies for describing ABMs, such as the ODD protocol [17] and its ODD+D extension [18] are a step forward to overcoming this claimed lack of model transparency. Finally, ABMS undoubtedly has more data processing and algorithmic computing requirements. However, this is an issue that should become less and less influential in the choice of ABM, given the tremendous advances in information and communications technologies (ICT), including cloud computing and multiprocessing.

The remaining of this deliverable D3.1 is organised as follows: section 2 presents the general description of AGRICORE model using the ODD+D protocol. Section 3 describes the execution workflow of the AGRICORE tool, including its different phases: initialisation, calibration and multiperiod dynamic simulation. Section 4 compiles the different equations that allow computing the dynamic evolution of the agents between one tick and the following one, as a consequence of their state and actions. Finally, section 5 presents the conclusions of this deliverable.

2 Description of AGRICORE Tool using ODD+D Protocol

The ODD+D (Overview, Design concepts and Details + human Decision-making) protocol [18] aims to describe agent-based models that include human decision-making in order to ease the comparison of different models and expand and refine the usual descriptions used. It facilitates a clear and comprehensive description of ABMs in a standardised way, with an emphasis on human decisions and which includes the empirical and theoretical foundations for the choice of decision model. ODD+D has three main sections (Overview, Design concepts and Details), each one of them with a set of subsections and topics, shown in Table 1.

Table 1: Structure of the ODD+D protocol

I. Overview	II. Design concepts	III. Details
I.1 Purpose • Purpose • Potential users	II.1 Theoretical and empirical background • Underlying concept, theories or hypotheses • Assumptions • Justification for chosen decision models • Empirical data foundation • Level of aggregation of empirical data	III.1 Implementation details • Tools used for model implementation • Public availability of the model
I.2 State variables and scales • Kinds of entities • Attributes of entities • Exogenous factors/drivers • Spatial dimension • Temporal and spatial resolution	II.2 Individual decision making • Subjects and objects of the decision making • Basic rationality guiding decision-making • Decision-making taskflow • Adaptation • Influence of social norms or cultural values • Influence of spatial aspects • Influence of temporal aspects • Modelling of uncertainty	III.2 Initialisation • Initial state • Inisialisation variability • Arbitrariness in selection of initial values
I.3 Process overview and scheduling • Sequence of simulation events	II.3 Learning • Individual learning • Collective learning	III.3 Input • Input from external sources
	II.4 Individual sensing • State variables (endogenous and exogenous) self-sensed • State variable sensed from other agents • Spatial scale of sensing • Modelling of information exchange mechanisms • Costs for cognition	III.4 Submodels • Description of submodels • Submodels' parameters, dimensions and reference values • Rationale for submodels selection, parameterisation and testing
	II.5 Individual prediction • Data used for making predictions • Internal models used for making predictions • Accuracy of predictions	
	II.6 Interaction • Direct and/or indirect interactions • Factors affecting interactions • Representation of communication during interactions • Coordination of imposed or emergent network	
	II.7 Collectives • Imposed or emergent aggregations of agents • Ways for representing collectives	

	II.8 Heterogeneity • State variables and/or processes diferring between agents • Existance of heterogeeous decision-making models	
	II.9 Stochasticity • Random or pseudo-random processes in the model.	
	II.10 Observation (incl. emergence) • Collected data from the ABM • Emerging outputs	

The description of the AGRICORE model using the ODD+D protocol is presented below:

2.1 Overview

2.1.1 Purpose

The purpose of this model is to analyse at the microscopic level the effect of the implementation of the different monetary instruments and agri-environmental schemes associated with the Common Agricultural Policy of the European Union.

In this sense, it allows both quantitative predictions to be obtained ex-ante to the implementation of agricultural measures and the corroboration of key performance indicators ex-post to their execution. AGRICORE, therefore, aims to become a fundamental tool for the design and evaluation of the impact of public agricultural policies.

The model is fundamentally designed so that its results can be used to support the process of designing public policies by policymakers, both at the Community level in the EC and at the national level in the respective Ministries of Agriculture of the MS.

However, its main users are likely to be modellers from the European Commission's own research centres (JRC) or scientists from private academic institutions dedicated to agro-economics. They produce reports and studies based on the results obtained with the tool, which ultimately guide or support the decisions of policymakers.

2.1.2 Entities, state variables and scale

The model mainly considers individual farms (generically referred to as Agricultural Holdings), characterised by their constituent elements, namely: physical or legal personality, socio-demographic characteristics of the person(s) owning the farm, socio-demographic characteristics of the person managing the farm, amount of land available and allocation of activities on the farm, own mechanisation capacity, own labour capacity, and financial statement.

Although the tool allows the generation of as many individual agents as there are real farms in a given area, for the analysis of certain measures, some types of farming may not be of interest to the modeller. However, they should still be considered in the simulation due to the effect of their interrelationships with the farms of interest under study. In these cases, it is also possible to create super-agents (archetypical farms representing a set of farms of a certain typology or a complete production sub-sector).

The markets for the exchange of productive factors (for the moment only land) and for the sale of output products from the AHs are also contemplated in the tool, not as agents but as external modules.

Additionally, the ecosystem of which the farms are a part can be represented by computing the effect on its biotic and abiotic factors of the actions of the agricultural holdings. However, so far, the ecosystem is not constituted as an agent in itself and its state (the mentioned effects) is not fed back into the AHs.

A detailed list of attributes is presented in Table 2.

The main external factors in AGRICORE are:

- The climatic conditions, which determine the performance of the different activities of the agents in each simulation step (agricultural year/season).
- The policy instruments active during the simulation, which constrain the agent's space of possible actions and modify his preferences for some activities over others.
- The interest rates that modify the agents' ability to afford and repay loans.
- Market prices for productive inputs and outputs.

At the time of initialisation, each agent is assigned the code of the NUTS 3 to which it belongs, as well as a georeferenced location in Latitude and Longitude. This location is used to determine which other agents constitute the agent's neighbourhood and can be influenced by its actions, as well as in which land market each agent participates.

Each simulation step represents one agricultural season. A priori, this could coincide with an agricultural year, but it could also be reduced in the case of use cases with more than one cropping season per year. In general, a simulation step occurs each time holding companies have to make new decisions on resource allocation to productive activities. In each simulation step, the productivities of each activity are also computed as a function of the given climatic conditions and the prices that each holding receives for its total production.

Talking about spatial resolution in AGRICORE is complicated because referring exclusively to the agents, the minimum distance between the locations of a pair of agents can be arbitrarily small.

However, to calculate the annual yield, the biophysical models associated with the tool need both weather conditions and soil quality data. If we take into account that the spatial resolution of meteorological data (MERRA database, SPEI database) is never smaller than 25x25km, we could consider this as the global spatial resolution of the whole tool.

Table 2: Attributes of interest of the objects forming each agent (Source: own elaboration).

Objects	Agricultural Holding Structure	Agricultural Holding Owner(s)	Agricultural Holding Manager	Parcel (optional)	Crop enterprises	Livestock enterprises	Output Products	Economic Financial Statement	Ecosystem
Parameters	Number of owners, probability of generational renewal, geographic location (coordinates, NUTS3, further granularity).	Gender, Grade of innovativeness, Risk aversion level.	Gender, Grade of innovativeness, Risk aversion level, education level.	Geographic location (centre), shape (coordinates of the polygon vertices), area, allowed uses, etc. Soil quality parameters	Type, regional cultivation standards, average regional yield.	Type, regional breeding standards.	Type	(re)investment propensity, size synergies, rate of interest, tax rates, WD_{min}	Soil properties: number of layers, layer thickness, max. bulk density, clay, sand, silt, organic carbon Soil types. Aquifers' quality
States	Economic size (FADN), Type of Farming (FADN), land structure (total area, parcels), livestock units, machinery capacity, regular workforce.	Age, probability of generational renewal	-	Current use (crop/livestock/mixed enterprise) Soil quality status	Status of permanent crops (age and health).	Livestock herd status (species and ages).	Stock levels	Assets, liabilities, equity. Solvency, Liquidity and Profitability indicators.	Soil properties: vol. water content, bulk density, nitrate levels, erosion level. Current classification of ecosystems that make the Agricultural Holdings located in them potential recipients of AES. (e.g. nitrates-polluted areas)
	Structural (LP): land ownership management (buy/sell), available capital, livestock			Chemical management: total amount of manure,	Production Technology (if more than one has been	Production Technology (if more than one has been	Production utilization (sales, farm use, farm	Investment, loans, withdrawals	

Agromanagement decisions	management (buy/sell), Workforce management, Machinery capacity management, quotas (milk, manure, etc.) **Agroeconomic (SP):** allocation of resources to enterprises, land management (rent/lease), contracted machinery,	-	-	ammonia amount, nitrate amount	considered in the model)	considered in the model)	consumption, changes in stock)		-
Disturbances	Land prices, production factors prices, output product prices, public policies.	-	-	Influence of weather conditions on soil status, external pollution.	Plagues, meteorological conditions	Unexpected deaths, meteorological conditions	Outputs from external agents (and imports)	Taxes, accountancy regulations.	Deviations of abiotic Factors (temperature, rainfall, etc), plagues, patogens, others.
Outputs	Socio-economic impact (labour, rent), environmental impact (land use, emissions, water intake, pollution)	-	-	-	Actual Yield	Actual Yield	Product Revenue	Cash flow, profit/loss, balance sheet.	Ecosystem services impact

2.1.3 Process overview and scheduling

The complete sequence of the AGRICORE tool workflow is presented in section 3-AGRICORE Tool Usage Workflow.

2.2 Design Concepts

2.2.1 Theoretical and empirical background

The hypothesis of the long-term model is that model-based predictive control (MPC) can be used as an algorithmic model of decision-making in the human brain [19]. This implies that, in general, two different models can co-exist:

- ❖ A detailed and complex (generally non-linear) one, based on the economic-biophysical equations that determine the functioning of a farm as a commercial company and as an agricultural production unit. This model is the one used to simulate, season by season, the effect of the actions decided by the Farm Manager.
- ❖ And another model, which can be less detailed and simpler (for example linearised around some break-even point), which is used for optimisation and which represents the (simplified) idea that the Farm Manager has of how the financial and agronomic dynamics of the agricultural holding works, i.e. the mental model that allows him to predict what will happen to the financial and agronomic states of the farm if he applies one or other management decisions.

Therefore, the decisions made by agents at the structural and financial level are inherently bounded rational, due to both cognitive limitations (simplified mental model) and imperfect information (presence of unknown a priori and uncontrollable external shocks).

There are also some ad-hoc behavioural rules in the model, especially in relation to succession in the agricultural business. It is assumed that, at a certain pre-set age, farm owners retire. If they have descendants, they will continue the operation of the farm as long as the average annual income earned by the farm is higher than the average wage in the secondary and tertiary sectors in the geographical area where the farm is located.

As the AGRICORE tool intends to be an exhaustive microscopic tool able to assess the impact of very disparate policy instruments, there might be a lack of information to initialise some agent's attributes or modules' elements. When this happens an 'information gap' is detected. AGRICORE has proposed (deliverables D1.7 and D1.8) a methodology to fill these information gaps through different kinds of Participatory Research activities. In the particular case of the three Use Cases contemplated in the AGRICORE proposal, three respective survey campaigns have been implemented in Andalusia (Spain), Greece and Poland to disclose information leading to the quantification of parameters such as risk aversion, grade of innovativeness, technology transition costs, etc.

The AGRICORE tool has extensive data requirements, the level of aggregation of which is very heterogeneous:

- The data used to construct the synthetic populations of agents come primarily from the FADN microdata, and are therefore a statistical sample at the level of individual agricultural holdings. However, some attributes are generated from socio-demographic data sources that may be at the level of municipalities (LAU2) or provinces (NUTS3).
- Data from Participatory Research are at the level of statistical sample of individual farms.
- Climate and soil data used by biophysical models and other impact assessment modules (IAMs) are usually aggregated at the level of a geo-defined grid, or in case of equivalent administrative unit resolution at NUTS3 or NUTS2 resolution.

2.2.2 Individual decision making

There are two types of subjects of decision-making in the AGRICORE tool: i) public policy-making institutions and ii) managers of European agricultural holdings at different geographical scales. The object of the decision of the formers are the different options of alternative policy instruments that can be implemented at different scales (European, national, regional), whose impact is to be analysed or predicted. These various policy instruments constitute external disturbances for the AH agents.

The object of the decision of the latter is whether to adhere to these aid instruments (or agri-environmental schemes), provided that their status (set of financial-agronomic states) allows them to do so.

Focusing specifically on the decision-making by AH agents, the cognitive process takes place on two clearly differentiated but interrelated conceptual-temporal planes (Figure 2). Agents take their decision at both levels by solving two different optimisation problems, posed by the respective objective functions and a set of constraints.

On the one hand, decisions affecting the holding structure itself (basically the size of the holding in terms of land operated or its capacity for self-supplying productive factors such as machinery or labour) are taken in the long term by assimilating the manager's strategy to a model-based predictive controller (MPC). At this structural level (long-term financial optimisation), agents have two explicit objectives, which are to maximise net profitability (the ratio of profits to equity) while keeping the solvency ratio within safety margins. Fundamentally, this translates into determining the appropriate size of the farm (utilised agricultural area) and its own resources (machinery and/or permanent workforce). Each agent prioritises one objective or the other depending on its behaviour profile (i.e., depending on attributes such as risk aversion). On the other hand, in each simulation interval, assuming that the AH structure for that season is established, each agent makes decisions on the allocation of production factors to productive activities using an optimisation calibrated by means of positive mathematical programming (PMP). At this agronomic level, the agents try to optimise the mix of agro-livestock activities in order to maximise the expected profit from the sale of the agricultural products generated during the season.

Both steps are repeated iteratively for each simulation interval, after updating the agent's financial and agronomic states based on the yield and market realisation of the immediately preceding simulation step, using the dynamic model explained in section 4-Dynamics of AGRICORE agents.

A detailed description of these two optimisation-based behavioural models is presented in Deliverable D3.2.

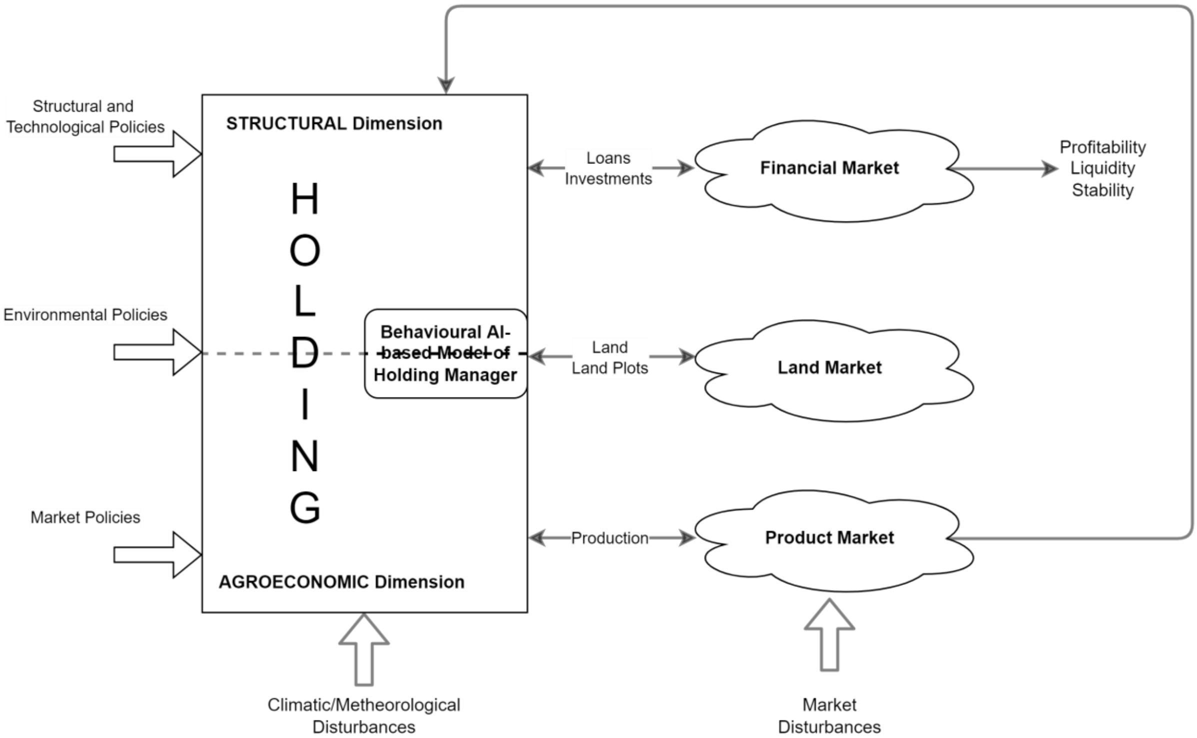

Figure 2: AGRICORE's agricultural holdings diagram concept

Agents do adapt to changing exogenous variables and/or external shocks. For example, within the decision of renting in/out. Based on the market prices, the agent will decide whether to rent or not. Likewise concerning buying/selling land. Another example is the modification of the decision-making due to agricultural policies or incentives (e.g., subsidies for ecological farming).

At the moment, neither social norms nor cultural values are covered by the AGRICORE tool. However, agents belonging to the same geographical vicinity may know the actions taken by their neighbours and (at least partially) the results of these actions, so that eventually terms could be added to the individual cost function in order to align the individual practice of each agent to the best practices observed at the local or regional level.

Spatial location indirectly influences the decision-making process of actors, as the quality of the soil that each actor owns (given by geo-referenced soil quality data) determines the average expected yield for different activities. Additionally, the simulated climatic conditions also vary according to the geographic location of the agents, indirectly influencing their future decisions.

In this initial version of AGRICORE, agents have no memory (i.e., they do not take into account the degree of performance of past actions when making the next decisions).

However, time does play a role in the optimisation process. On the one hand, as MPC is a sliding horizon technique, it requires a dynamic prediction model. On the other hand, in the objective function itself, and depending on the risk aversion profile of the agent, immediate profits or the expectation of higher future profits can be prioritised.

The different level of risk aversion and the different levels of propensity to innovate are taken into account by the respective coefficients that weight the terms of the objective function, representing, for example, a different penalty for the same deviation of the actual solvency level from the target solvency level. Beyond that, the model does not currently include other mechanisms for incorporating uncertainty into the decision process.

However, the very nature of the Economic-MPC makes its *stochastication* relatively straightforward by including alternative future scenarios and weighted optimisation according to the probability level associated to each scenario.

2.2.3 Learning

Learning has not been considered so far in the AGRICORE tool, partly because the initial period foreseen for the multi-year simulation of the use cases is $N_h = 7$years, which coincides with the usual period of validity of each EU CAP reform. This timespan is too short for the agent to be able to draw the consequences of its (non) adherence to a given policy instrument and change its actions accordingly. For similar reasons, collective learning has not been considered in the tool for the time being either, beyond the possibility of agents imitating best practices at the local level.

2.2.4 Individual sensing

Agents (Agricultural Holdings) perceive the average cost of land exchange from the Land Market Module (LMM) and use it for defining their next Land trading intentions.

They also exogenously receive (or endogenously generate) output price prediction sequences.

Agents exchange information both directly and indirectly.

Through the Land Market Module, agents participating in the same local market house (i.e. agents belonging to the same municipality or geographical vicinity) can "hear" to the offers made by other agents for certain plots of land, thus being able to intuit what their unit land valuation is.

Additionally, agents exchange technological information at regional level, by including the concept of "frontier farm" in the short-term model, as explained in deliverable D3.3.

The mechanisms for obtaining information from other agents have not been explicitly modelled, beyond the Land Market Module's own architecture, based on discrete auctions with public bids (which are known to all participants, who can adapt their future bids in response).

However, no models of diffusion mechanisms for innovations or best practices have been explicitly incorporated so far. Therefore, costs associated to information gathering or cognition are not considered either.

2.2.5 Individual prediction

As agricultural managers do in real life, agents use exogenous forecasts from the relevant institutions to predict future prices and weather conditions.

At the financial-structural level, agents rely on the MPC prediction model, which is nothing more than a simplified (possibly linearised) version of the non-linear dynamic model given by the equations in section 4-Dynamics of AGRICORE agents.

At the agroeconomic-agricultural level agents maximize gross income, subject to the structural and policy constraints specific to their farm, but also bounded by the possibility of interacting, by exchanging production factors, with other farms in the same region.

Prediction errors are inherently covered by the use of the MPC itself. Deviations between the predicted state and the state actually achieved can be due to both differences between the prediction model and the simulation model (representing the limited cognitive capacity of the AH managers) and external disturbances applied during the simulation of each cropping season (representing the effect of uncertainties that the AH cannot accurately predict nor of course control).

2.2.6 Interactions

Both types of interactions are used in the tool. Direct interactions among agents occur in the Land Market Module while buying/selling land before each agricultural season. Indirect interactions occur in the exchange of technology through the use of the hypothetical frontier farm (D3.3).

The possibility for an agent to interact with other agents depends fundamentally on the geographical location of all of them, either because they belong to the same administrative area (municipality, agricultural district, etc.) or because they are all within a certain range of physical distance.
In principle, these interactions are not limited by the typology of agents, and exchanges of land and technology are possible between any pair of agents. The interactions that take place between agents within the Land Market Module are modelled as explicit bid or ask messages that are broadcasted to all agents belonging to the same local market.

There are no coordination networks between actors that modify their community behaviour.

2.2.7 Collectives

Agents do not belong to or form any collectives.

2.2.8 Heterogeneity

The heterogeneity of the actors is one of the main characteristics of the AGRICORE tool. During initialisation, in the process of generating the synthetic population, each agent is endowed with its own values for its attributes, while ensuring that the statistical distribution of the values of these attributes reproduces up to a certain minimum threshold of adjustment the statistical distribution of the values of these attributes in the real population.

The main states that distinguish one agent from another are the economic dimension (given by the value of their gross output), their physical size (given by the amount of land they own), and the type of farming (given by the distribution of land to agricultural activities).

All actors have the same objects of decision: the development of their structure (own productive capacity) and the optimal allocation of these resources to specific agricultural activities. Although all agents share the same structure of optimisation problems that define their behaviour, some aspects of these problems may vary between agents:

- The size of the agents (their economic size) may determine different financial capacity (e.g. with respect to the maximum degree of indebtedness, or with respect to the target solvency value).
- A policy may be applicable only for a subset of agents that meet a set of requirements (whose states are between a certain range of values). In this case, the objective function or constraints defining the optimisation problem will be modified only for that subset of agents.

2.2.9 Stochasticity

Subsequently, during the simulation, the weather conditions observed during each season can be set deterministically (this will be the case for ex-post impact analyses where historical weather observation data will already be available) or they can be generated randomly (this will be the case for ex-post impact analyses where no observations exist because of future seasons).

2.2.10 Observation

Data on the actions taken by each agent (land acquisition and alienation, loan taking and repayment, selection of productive activities) and on the effects of these actions on their states are stored after each simulation interval, equivalent to a full agricultural season. This allows their partial visualisation during the simulation itself.

However, the full usefulness of the collected data is not obtained until the simulation of the multi-year period is completed. It is then when, with all the data for each season, the analysis of the dynamic variation of the financial and agronomic statements at the individual and sector average level can be made. The aggregated data can also be used to calculate the socio-economic, environmental and ecosystem service impact through the corresponding Impact Assessment Modules.

The key results that emerge from the full use case simulation are the level of adherence of the agents to the policy measure(s) under analysis, and furthermore, the effect that these policies have on the financial and socio-economic states of each agent, with respect to the baseline scenario which would be the maintenance of the existing policies at t=0 and/or the implementation of no additional policy modifications.

At the global level, the effect that the variation in agricultural activities produced by the tested policies has on the environment (variation in CO_2 emissions, water use and pollution) and ecosystem services (variation in the area of habitats suitable for animal species, variation in animal populations, complementary activities to agriculture, etc.) emerges.

2.3 Details

2.3.1 Implementation details

The main programming language used to implement the different modules and to exchange information among them is Python 3. The 'R' language is also used for the extraction of Bayesian Networks within the SPG process. In addition, for the mathematical modelling of the ST optimisation, GAMS is used as programming language and IDE. For solving the optimisation problems faced by the agents, external solvers such as Gurobi and MOSEK are interfaced.

All the project deliverables and technical documents will be available at the project's website. The code is open-sourced at the project's GitLab public repository.

Part of the generated data is available at the AGRICORE Zenodo open-data repository.

2.3.2 Initialisation

At time t=0 of the simulation, the synthetic population of agents reproduces at a global level the real population on which the effect of the policies wants to be assessed. That is, the probability density distribution of the value of each attribute among the agents of the synthetic population reproduces (up to a certain level of adjustment) the probability density distribution of that attribute in the individuals of the real population.

This is achieved by using Bayesian networks for the sequential generation of values and for the conditional generation of values for those attributes that correlate with each other (see deliverables D2.3 and D2.4).

A priori, the composition and values of a synthetic population should not vary between two different simulations. This would not make sense as it would make it difficult to determine whether the different impact observed when simulating two different policies is due to the difference between the policies themselves or to the baseline difference between the initial populations of agents.

What can vary are the weather conditions that will be used to simulate each season (and on which the yield and therefore the economic performance of the agents will depend). These conditions can be initialised deterministically or pseudo-randomly and can be different between two consecutive simulations of the same policy measure (e.g., to see the level of resilience of the agents adhering to the policy instrument under different climate scenarios).

2.3.3 Input data

In general, it can be stated that all external disturbances affecting the realisation of each agricultural year within the iterative multi-annual simulation come from external data sources and/or models.

On the one hand, the weather conditions that are used to simulate the yield of each activity on each farm come either from databases with historical observations or from predictions constructed using external climate indicators and models.

On the other hand, the yield simulation itself is done using external biophysical models (e.g., DNCN, Wofost, STICS) through AGRICORE connectors built for this purpose. The enumeration and details of these biophysical-climatic data sources and models are presented in deliverable D3.4.

There are also external modules that affect the dynamic behaviour of AH. They are more related to the agents interacting individually or collectively during each simulation step to calculate their control actions and the partial effects of these actions.

- **Land Market Module**: decisions to lease/rent land are the outcome of the short-term economic results (positive gross margin); decisions to buy/sell land are the result of long-term economic optimisation/financial ratios. However, with respect to these latter decisions, one thing is the agent's intention, and another is whether this intention can be realized (e.g., a desire to buy is only realised if there is a desire to sell and the offered and asked prices are compatible). Interactions for land transfer are simulated through the Land Market Module (D5.2). This module allows farmers to place offers (bid/ask orders) according to prevailing land market prices. These offered land prices might be a function of geographical location, land cover, land quality and allowed use(s) of the plot(s).
- **Product Market Module:** This module simulates the dynamics of output market prices based on the aggregated demand (exogenous parameter) and on the aggregated supply produced by the agents (internal supply) or by other actors (such as the external supply due to importations). Other input markets could be considered (above and beyond land), such as labour, manure, fodder and young animals.

In addition, the policy environment module (D5.7) provides the representation of the legislative instruments associated with the public policies whose impact is to be evaluated. This representation consists of a series of mathematical formulations that modify the rules (optimisation problems - D3.2) that determine the behaviour of the agents. In the case that these instruments are not constant (e.g., a subsidy decreasing over time), the sequence of monetary amounts of this instrument is a time series of values incorporated externally into the simulation.

2.3.4 Submodels

The submodels used to dynamically update the state of each agent are presented in section Dynamics of AGRICORE agents4.

3 AGRICORE Tool Usage Workflow

AGRICORE's agent-based simulation (ABS) workflow can be divided into three phases, namely, i) Initialisation, ii) Calibration and iii) Multi-year simulation. Once the multi-year simulation is finished the results (set of time series with the evolution of agents' states, inputs and outputs) are stored in a file. The iv) computation of adequate KPIs using the Impact Assessment Modules and the v) Visualisation of results are optional phases which will be normally performed but not mandatory.
Each phase is composed of several stages executed sequentially. In the first stage, the agents' technology alternatives and weather conditions are initialised based on their location. In the second stage, the model agents' parameters are calibrated in order to properly simulate their decision-making strategies (e.g. PMP calibration of the agroeconomic model). The third stage consists of simulating, on a yearly basis, the agents' decision-making and its consequences on the dynamic evolution of the agents' states. A multi-year simulation is completed through the iterative realisation of individual agricultural years/seasons. These individual realisations mainly consist of simulating the yields and prices for each activity and each agent in the population, given certain climatic conditions during the simulated season. The stages forming the Initialization and Calibration phases are executed only once at the beginning; meanwhile, the stages of the multi-year simulation are recursively executed as many times as the number of years in the selected period (which must be relevant for the detection of impacts of the assessed policies)

Figure 3 schematically depicts the execution workflow of an AGRICORE use case. Moreover, further details for each stage and its associated subprocesses are given in the following subsections, following their execution order.

3.1 Initialisation

The objective of the initialisation phase is to extract and/or generate all the data necessary for the calibration and/or simulation of the agents. This includes the different technologies assumed to be available for the operation of each agro-livestock activity, the different performance (productivity) of each technology-activity pair in relation to different typical annual climatic conditions, and the actual climatic-meteorological conditions under which the performance of the agents will be simulated in each of the seasons of the multi-annual period of the simulation (Figure 4).

In the field of agricultural economics, the term 'technology' encompasses all the techniques applied (including the equipment and/or machinery used) to control the growth and harvesting of animal and plant products. Therefore, in planning each agricultural year, an AH manager must choose not only the activities to be carried out but also the technology with which to operate them. At the planning level, the same crop operated with two different technologies means two different and alternative activities (e.g., wheat cultivation with conventional practices vs. wheat cultivation following organic practices), each with different expected productivity, expected costs and expected income.

Within the initialisation stage, the first objective is to determine the set of agricultural technologies available to the agents. The biophysical modelling of technical alternatives (see deliverable D3.4) considers two categories of practices relating to conventional and organic agriculture. In addition to conventional and organic, they are further categorised by two different irrigation levels, three different fertilization levels, and two different tillage practice possibilities. Only non-tillage methods are taken into consideration for organic technology. Finally, all feasible combinations of techniques result in 30 different technologies. The expected yield for each activity and each technology is extracted at the NUTS3 level using external biophysical models.

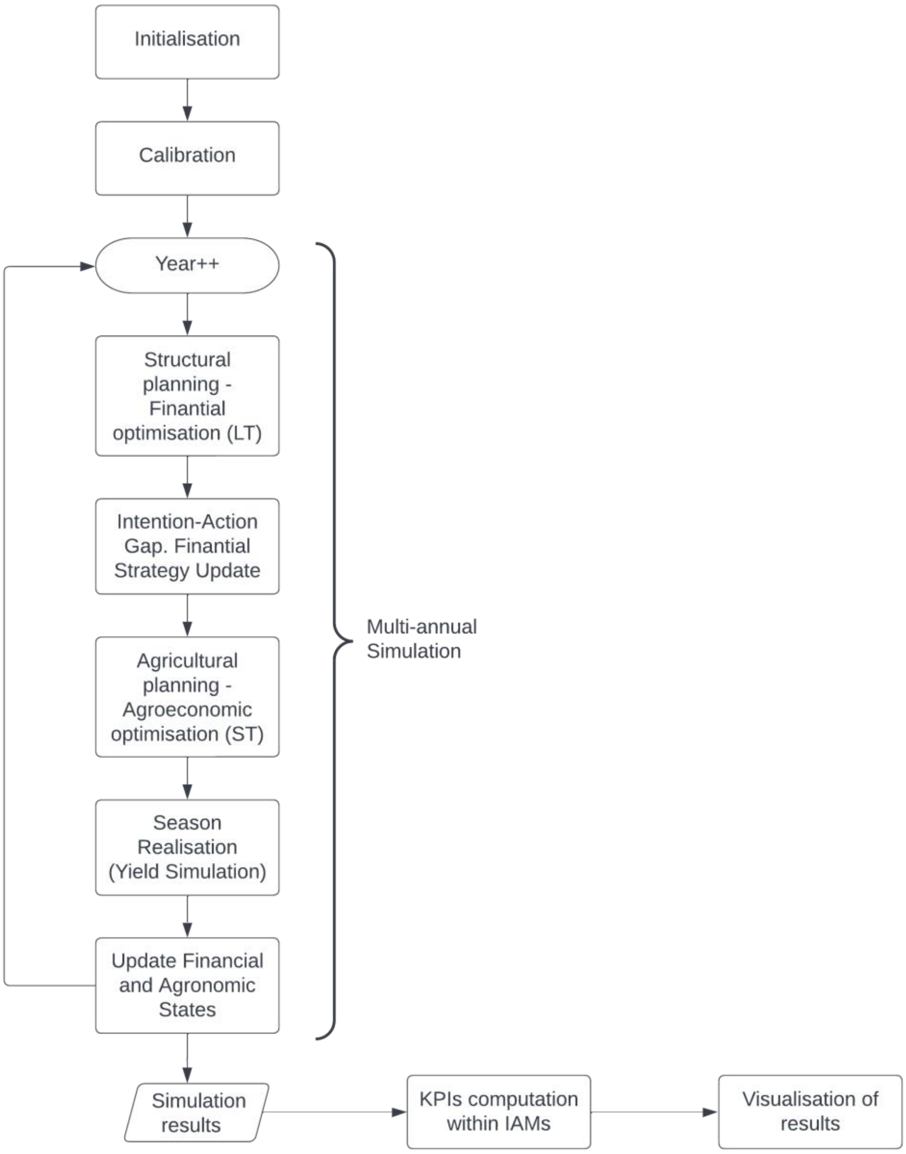

Figure 3: AGRICORE Tool Workflow

The second stage of initialisation consists of determining the meteorological conditions to which the agents will be exposed during the simulation, depending on the NUTS3 in which each of them is located. If the use case is ex-post, accurate weather records may exist for all years under study, which can be downloaded from an external weather database.

If the use case is ex-ante, or if one wants to simulate what-if cases with different weather conditions than those that actually occurred, it is necessary to generate them. This is done in two steps. First, the type of weather year (standard, wetter than normal or less wet than normal) is chosen (randomly or user-selected). Once chosen, the daily weather records for a past year of the same type as the one selected are extracted from an external database. For the classification of climatic year types and the selection of the individual year type, use is made of the Standardised Precipitation-Evapotranspiration Index (**SPEI)**, which is a multiscale drought index based on climatic data. The detailed methodology and data processing performed in the climate module for this stage are presented in D3.4.

Finally, the average productivity of each activity-technology pair in the different types of climatic year is determined at NUTS3 level. This information is important because it will be used for calibration (and optimisation prior to each of the seasons of the multi-annual simulation), but also because it can be used as the yield actually occurred if, during the season Realisation stage (see Figure 5), external biophysical models cannot be used for yield simulation.

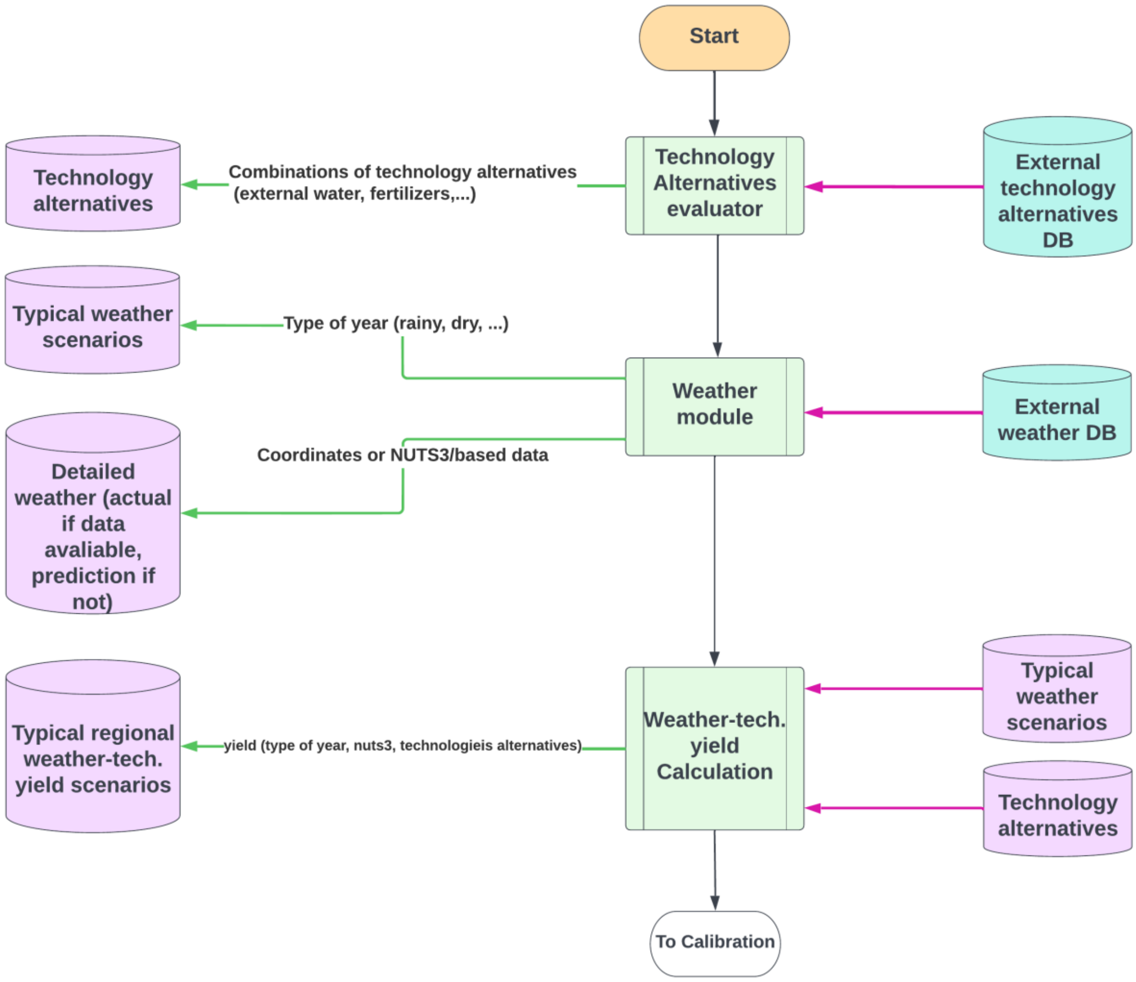

Figure 4: Detailed workflow of Initialisation stages

3.2 Calibration

The main objective of the calibration phase is to ensure that the model is able to reproduce into the synthetic population the allocation of activities actually observed in the field for the base simulation year. Actually, this baseline scenario is already given in the synthetic population by the generation process itself (i.e., the sum of the activities assigned to the agents coincides with the real areas and productions observed for the base year). What the calibration ensures is that, even if all agents were to perform an initial optimisation procedure, the resulting allocation of activities would match the initial allocation given for the SP in the base year (i.e. the sum of the activities allocated by the agents coincides with the actual allocation observed for the base year). This scenario constitutes the baseline for the simulation iteration (the first hypothetical agricultural year) to which the following agricultural years will follow in an iterative logic, integrating long-term information into the short-term module and vice versa.

The calibration process computes the explicit and implicit costs for each activity, resulting in a symmetric, positive semi-definitive matrix ($\frac{1}{2}\boldsymbol{xQx'}$), used within the agro-economic optimisation.

To perform the calibration procedure, the input, from the synthetic population generation, listed below are required:

- the vector of realized output prices, expressed in € per unit of product: $\boldsymbol{p_n}$;
- the vector of realized production level expressed in ha per crop: $\boldsymbol{x}^*$;
- the total known variable costs aggregated for all activities at farm level TVC;
- and the coupled payments (in $€/ha$).

The matrix $\boldsymbol{Q}$ generated as a calibration output, in addition to the explicit and implicit cost per crop, describes the substitution/complementarity ratios among all the cultivated crops in the area, as well as the increase of marginal cost per unit of product for each crop, at regional level. Furthermore, the calibration returns the vector of marginal cost deviations per farm (the distance between individual farms' cost functions and frontier cost function).

The calibration phase, integrated into the Positive Mathematical Programming approach used in this project is described in detail in deliverable D3.2.

3.3 Simulation

The aim of the simulation phase is to reproduce the dynamic evolution of the agents (including their states, decisions and outputs) and their environment during the real period represented by the simulation period, in the presence of the public policies whose impact is to be evaluated. As reflected in Figure 5, the simulation phase is composed of several stages. These stages are executed iteratively (each tick representing an agricultural year/season) until the end of the selected simulation period.

3.3.1 Structural planning (Financial optimisation)

Within the simulation of the agent's decision-making, the first step is financial-based structural planning. Financial decisions are often medium- and long-term decisions related to actions such as investments or loans. Therefore, the optimisation is performed for a long-term horizon and not just for the upcoming agricultural season. It is important to recall that AGRICORE is not a farm-advisory system, so "optimisation" here does not mean finding the best possible solution (as seen by an omniscient observer external to the simulation) but reproducing the decision the agent would take according to its criteria. Consequently, the output of the process will be the agent's planning at the financial level aiming to optimise profitability and solvency.

The process receives as inputs several variables which can be categorised into three groups:

- <u>Intrinsic characteristics of the agent.</u> This group enclose all the variables that define the agent and distinguish it from the others. The variables included are the age and gender of the agricultural holding owner, heirs, risk aversion indicator and innovation factor. How those variables shape the behaviour of each agent is detailed in D3.2.
- <u>Exogenous variables.</u> This group enclose all the external variables such as the crops' average price and yield, depreciation of actives and interest on loans.

- Financial previous state. This group enclose all the variables to be found on a balance sheet of an agricultural holding. Therefore, the variables included are owned land, available land (including rent land), mechanisation capacity, short-/long-term liabilities, deposits, equity, farm net income, gross farm income, liquidity ratio, solvency ratio and net profitability ratio. In the first iteration, the values of those variables are those assigned in the generation of the synthetic population. As the simulation goes on, the states are updated based on the decisions made by each agent and their consequences.

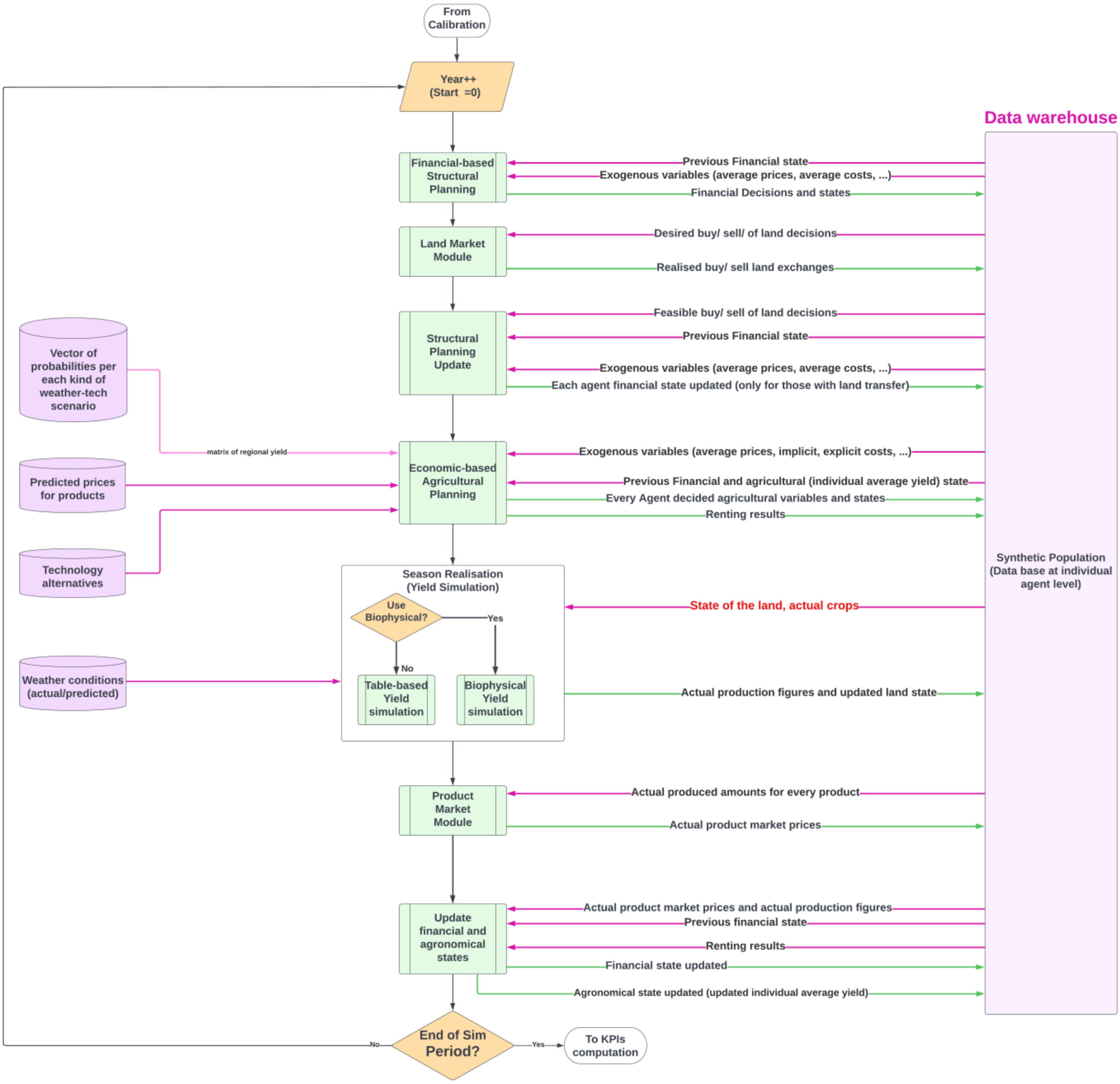

Figure 5: Detailed workflow of the Simulation Phase

The dynamics of the financial model of the agent (detailed in the next section), constrain its planned decision strategy. The chosen planning determines if the agent intends to buy or sell land, increase or decrease its mechanisation capacity, and the acquisition of new loans required to finance the above investments beyond the use of own funds on deposits.

3.3.2 Resolution of Intentions vs. Actions Gap and Update of Financial Strategy

The intentions of the agents are determined through financial optimisation as explained above. However, those intentions might not necessarily be fulfilled, as they may collide or be incompatible with the intentions of other agents. For example, an agent's intention to buy a quantity of land cannot be realised if there is no other agent intending to sell land or if, if there is one, the offers to buy and sell are not compatible. Otherwise, the original intention of buying land is not realised and the structural planning might need to be re-evaluated. This is called the 'intention vs. action' gap. Land exchanges are resolved through an auction system in the land market module (LMM), presented in D5.2. The resolution of the LMM defines whether the agents can follow their initial structural planning, or they should recompute it considering the result of the land market. In any case, the financial status must be updated accordingly.

3.3.3 Agricultural Planning (Agroeconomic optimisation)

The agroeconomic-based agricultural planning is performed in the short-term period (for the following agricultural year) and it is subject to structural, financial and policy constraints.

The objective of this phase is to replicate the manager's decision-making process related to the agricultural operation of the AH as a productive system (i.e. the allocation of resources to activities). The required inputs are:

- Structural constraints: defined by variables coming from the long-term financial strategy, more specifically: available land and the total amount of available capital.
- Exogenous variables: included in this category are the price expected for each product and the average rental price per hectare.
- Agricultural holding initial state: land allocation per activity at t_0, defined in the synthetic population generation. The initial land allocation is consequently updated, year by year, based on the results of the previous season and the subsequent optimization runs in t_k.
- Technological AH initial state: these are the agricultural management practices currently adopted by the AH's owner and the set of available technology alternatives[1], that the AH's owner can chose as reaction to policies interventions, subsidies and market shocks.
- Biophysical data: information coming from the biophysical module that might impact the yield of the selectable activities.

The output of the process is the new land endowment (rent-in/-out), the land allocation per crop/livestock and the technologies implemented.

3.3.4 Realisation of immediate agricultural season (Yield simulation and Product price determination)

At this stage, the agro-economic outcome of the immediate agricultural season is reproduced for each agent, which will generally be different from the expected outcome used to carry out the structural and agricultural planning.

The realisation of the agricultural season includes:

1. Simulating the yield obtained by each agent in each activity. This is normally done through the use of external biophysical models, taking into account the type of technology used for cultivation/breeding and the climatic conditions experienced (determined during the

[1] This set remains invariant for the whole multi-annual simulation period (i.e. technology improvements or emergent technologies are not contemplated)

initialisation phase). Only in case no connectivity with external biophysical models is available, the yields tabulated during the initialisation phase for each technology and each type of climatic year can be assumed as realised yields. Details on both alternatives are provided in Deliverable D3.4.

2. Simulating the price received by the agents for each unit of product. After the previous point, each agent has the total quantities of each product associated with the activities it has operated during the year. After the eventual reduction of own consumption and farm use, the remaining quantity is assumed as production supplied to the market. The individual productions of all the agents are aggregated to calculate the global production by product and geographical region. This overall production is the input to the product market module (D5.3), which returns the average price per product and region.

3.3.5 Update of Agricultural Holding states (Demographic, Financial and Agroeconomic)

The last step of the simulation process, before starting the iteration for the next campaign/year, consists of updating all the agent farm states. First, the agronomic state of the agents, such as the state of their agricultural land as a result of the implemented activities and technologies, is updated. For this, use is made of the results of the same biophysical models used for the yield simulation, which also provide the effect of the agricultural activities on the soil.

Secondly, the financial statements of each farm are updated, including also the results derived from land rental/lease. For this purpose, the revenue resulting from their farming operation (yield*price) and the costs associated with this production are used. After deducting the amount allocated to family support and the amortisation of liabilities, the farm provides a new set of financial statements (liquidity, solvency and net profitability).

Although obvious, it is necessary to mention the need to update the age of the managers and owners of each Agricultural Holding, as this attribute influences the strategic planning of the AH agents.
Finally, some global agronomic statements, such as land allocation or regional average yields, may need to be updated.

The equations defining the dynamic financial model that allows the updating of the agent's states are presented in the next section.

4 Dynamics of AGRICORE agents

Agricultural Holding managers' decisions have an effect not only on the immediate season but also on subsequent seasons. On the one hand, actions taken at the structural level (acquisition or disposal of land and/or machinery) require financing that may exceed the AH's own funds at the time, making it necessary to apply for loans. These loans increase the company's liabilities, impose a series of future repayments and may limit the ability to obtain liquidity in the short term. Also the net profitability of the agricultural holding company, as a trading company, determines the willingness of the farm owner (and his family) to continue farming or to cease operations and move to a non-primary sector. On the other hand, purely agricultural decisions (choice of activities and technologies) also have an effect on the productive infrastructure (land and/or livestock) of the agricultural holding. For example, agricultural experience recommends certain types of annual crop rotations that preserve the nutritional properties of the soil, avoiding soil exhaustion. Therefore, the past and present agronomic decisions of farm managers have an effect on their potential future productivity, even if they are not (fully) aware of it.

4.1 Dynamic evolution model of the financial status of an AGRICORE (Agricultural Holding) agent.

The objective is to create a model that mimics the time evolution of the Financial Statements (Balance sheet and Profit & Losses statement) of an AH based on its previous financial status and the decisions taken by its manager. To that end, each one of the portfolios that impact the Balance Sheet of the company must be addressed and its time evolution modelled.

Let $t \in \mathbb{R}^+$ be the time (time is discretised into agricultural seasons).

Let $TA(t)$ be the set of Total Assets that a company owes at time t:

$$TA(t) = FA(t) \cup CA(t) \tag{1}$$

where $FA(t)$ refers to the set of fixed assets at time t and where $CA(t)$ refers to the set of current assets at time t.

It is assumed that:

$$FA(t) \cap CA(t) = \emptyset, t \in \{1, \dots, T\} \tag{2}$$

Let's consider, without loss of generality, that the only fixed assets a farm company can have are lands and machinery. In this sense, the set of fixed assets is composed by:

$$FA(t) = A(t) \cup M(t) \tag{3}$$

where $A(t)$ denotes the set of owned land and $M(t)$ denotes the set of owned machinery at time$t$. Let $A_t \in \mathbb{R}^+$ be the total value of all owned land at time t. It can be assumed also that the only action that can vary the valuation of owned land are buy/sell actions. In this sense, A_t evolves as follows:

$$A_{t+1} = A_t + B_t^L \quad (4)$$

where $B_t^L \in \mathbb{R}$ refers to the buy/sell decisions (initially they are buying/selling desires that may be realised in the Land Market or not). $B_t^L > 0$ then implies the desire to increase the value of owned land (and therefore the amount of owned land) and represents a buying intention; $B_t^L < 0$ then implies the desire to reduce the value of owned land (and therefore the quantity of assets) and represents a selling intention.

On the other hand, let $M_t \in \mathbb{R}^+$be the value of owned machinery at time$t$. Again, we can assume without loss of generality, that the only actions that can vary the quantity of owned machinery are buy/sell management decisions. In this sense, M_t evolves as follows:

$$M_{t+1} = (1 - dep_t)M_t + B_t^M \quad (5)$$

where $dep_t \in (0,1)$ refers to the depreciation of machinery with time and $B_t^M \in \mathbb{R}$ refers to the buy/sell actions using the same sign criterion as for B_t^L.

Let $FA_t \in \mathbb{R}^+$ be the sum of the values of all the elements of $FA(t)$. According to the expressions (3), (4) and (5), the temporal evolution of the value of the Fixed Assets of the holding can be modelled as follows:

$$FA_{t+1} = A_{t+1} + M_{t+1} = A_t + B_t^L + (1 - dep_t)M_t + B_t^M \quad (6)$$

Let $CA_t \in \mathbb{R}^+$ be the value of all the elements in $CA(t)$. Also, let's consider that the only possible way for the holding to convert assets to cash is through money saved in deposits ($D_t \in \mathbb{R}^+$):

$$CA_t = D_t \quad (7)$$

Then, the evolution of D_t in time is as follows:

$$D_{t+1} = max(0, (1 + r_t) \cdot D_t + FNI_t - N_{fm} \cdot MEPI_t) \quad (8)$$

where $r_t \in \mathbb{R}^+$ refers to the interest rate during interval$t$, $FNI_t \in \mathbb{R}^+$ is the Farm Net Income for season$t$, N_{fm} is the number of members of the farm owner's family, and $MEPI \in \mathbb{R}^+$ refers to the Multiple Effects Public Income Indicator (the minimum annual living income considered as fair by the national government for year t). Notice that the following assumption has been made:

$$FNI_t - N_{fm} \cdot MEPI_t \overset{\text{is deposited in}}{\rightarrow} D_{t+1}, \quad (9)$$

i.e., all the margin of the Farm net Income after withdrawing what the farmer's family needs to live is assumed to be saved as bank deposits and will be available for the holding (and also for the family) in the following agricultural season.

Let $LT(t)$ be the set of all the long-term debt at time$t$, and let's define $LT_t \in \mathbb{R}^+$ as the sum of the values of long-term loans at time$t$. Also, let $m_L \in \mathbb{N}$ be the average maturity of the loans historically acquired by a holding. The evolution of LT_t in time can be modelled as follows:

$$LT_{t+1} = (1 - \frac{1}{m_L})LT_t + L_{t+1} \tag{10}$$

where $L_{t+1} \in \mathbb{R}^+$ represents the management decision of acquiring a new amount of lent money. To model the behaviour of long-term liabilities in this way, we need to assume that the loans are not paid off after the maturity term and, instead, for finite times there will always be a residual part to be paid off. Find the rationale below.

T is the maximum time of forecast (prediction horizon), so for $t > T, L_t = 0$. For that timesteps let $n \in \mathbb{N}$. Then,

$$LT_{t+1} = (1 - \frac{1}{m_L})LT_t \Rightarrow LT_{t+n} = (1 - \frac{1}{m_L})^n LT_t > 0 \tag{11}$$

So, for finite times we have:

$$LT_t > 0. \tag{12}$$

On the other hand, when $n \to \infty$:

$$0 < (1 - \frac{1}{m_L}) < 1 \Rightarrow \lim_{n\to\infty} (1 - \frac{1}{m_L})^n = 0 \tag{13}$$

This yields to:

$$\lim_{t\to\infty} LT_t = 0 \tag{14}$$

Long-term liabilities are used to pay for land and machinery acquisitions.

According to all previously exposed, the Equity of the holding at time t (E_t) evolves as follows:

$$E_{t+1} = FA_{t+1} + CA_{t+1} - LT_{t+1}. \tag{15}$$

The total output from crops and products (TC_t) of a farm in a specific season step can be modelled as follows:

$$TC_t = PYLD_t \cdot A_t, \tag{16}$$

where $PYLD_t \in \mathbb{R}^+$ refers to the average ratio of value of production per unit of land.

The total specific costs (SC_t) can be modelled as follows:

$$SC_t \approx C_{avg,t} \cdot A_t, \tag{17}$$

where $C_{avg,t} \in \mathbb{R}^+$ is the average ratio of price cropping activity of the farm per unit in the value of land.

In this sense, the Gross Farm Income produced at time t (GFI_t) can be modelled as follows:

$$GFI_t = TC_t - SC_t - c_{ov} = (P_t \cdot YLD_t - C_t) \cdot A_t - c_{ov}, \tag{18}$$

where $c_{ov} \in \mathbb{R}^+$ are the farming overhead costs.

On the other hand, Farm Net Income obtained at time t (FNI_t) can be mathematically expressed as follows:

$$\begin{aligned} FNI_{t+1} = \quad & GFI_{t+1} - dep_t \cdot M_t + SbT_{t+1} - VAT_{t+1} - EF_{t+1} \\ & - \frac{1}{m_L} LT_t - (B_t^L + B_t^M) \\ = \quad & (P_t \cdot YLD_t - C_t) \cdot A_t - dep_t \cdot M_t + SbT_{t+1} - VAT_{t+1} \\ & - EF_{t+1} - \frac{1}{m_L} LT_t - (B_t^L + B_t^M) \end{aligned} \tag{19}$$

where $SbT_t \in \mathbb{R}$ refers to the subsidies and taxes on investments, $EF_t \in \mathbb{R}$ accounts for external factors and $VAT_t \in \mathbb{R}$ refers to the value added tax.

The equations (1)-(19) presented above represent the financial dynamic model of the agricultural holding, to which the optimization problem in deliverable D3.2 is subject. Let $\boldsymbol{\mathcal{M}_f}$ designates this model. At the beginning of the iteration corresponding to each of the agricultural campaigns, during the structural planning stage, each agent determines its financial actions $u_f(t) = \{B^L(t), B^M(t), L(t)\}$ by solving the optimisation problem derived through Model Predictive Control (MPC) presented in the following deliverable D3.2.

4.2 Dynamic evolution model of the agronomic status of an AGRICORE (Agricultural Holding) agent.

It is not within the scope of the AGRICORE Project to model the complex non-linear chemo-metabolic equations that relate the growth of the organic mass of each crop with respect to the application of certain internal (e.g. application of one or another amount of fertiliser) and exogenous (e.g. different levels of precipitation) conditions.

Precisely for this reason, this task is 'outsourced' by using existing biophysical models with a wide range of applications, namely: DNDC [20], STICS [21], and WOFOST [22]. It is important to note that each of these three models uses a different approach to calculate yield growth and derived plant and soil effects, as explained in deliverable D3.4. It is also important to note that each model covers a different set of agricultural activities, listed in D6.3. The three selected biophysical models can calculate the crop biomass, yield, gas emissions from the cultivated fields, changes in soil status in each season for cropping activities. However, none of them is the livestock production model. In the DNDC model grazing is applied to grassland or pasture. as a factor influencing the grass yield. Grazing practice in DNDC is defined by specifying the livestock type, heads, and grazing duration. The data about livestock should be delivered to the model as input. and is used to quantify the feeding intensity and waste deposition of the livestock when they stay in the field. The aim is to calculate the actual yield of the grass in the fields. The livestock production is not calculated by these models as well as the actual livestock status. It may therefore be the case that several calls to different biophysical models may be necessary to calculate the agronomic evolution of the infrastructure (soil and livestock) of an agent.

Therefore, being fully generic and encompassing the different biophysically-based agronomical models as $\boldsymbol{\mathcal{M}_a}$, it is possible to compute the dynamics of soil states $S(t+1)$ and livestock states $Lv(t+1)$ as:

$$S(t+1) = \boldsymbol{\mathcal{M}_a}(S(t), u_a(t), d_a(t)) \quad (20)$$

$$Lv(t+1) = \boldsymbol{\mathcal{M}_a}(Lv(t), u_a(t), d_a(t)) \quad (21)$$

where $S(t)$ and $Lv(t)$ are the current Soil and Livestock states, $u_a(t)$ are the agronomic management decisions taken for the time period between t and $t+1$, and $d_a(t)$ are the external agronomic disturbances (climatic conditions). The effects of other external agronomical disturbances such as diseases and pests are not covered by the current versions of the selected biophysical models, but they could appear in future versions or be also incorporated using additional external models.

At the beginning of the iteration corresponding to each of the agricultural campaigns, during the agronomic planning stage, each agent determines its agronomic actions $u_a(t)$ by solving the optimisation problem derived through positive mathematical programming (PMP) presented in the following deliverable D3.2.

5 Conclusions

This deliverable presents the main features of the AGRICORE tool and ABM. It is a model that tries to reproduce what the behaviour of the agents (agricultural holdings) would be in the presence of a set of public policies, whose composition and parameters can be varied to test the different responses of the agents. The behaviour of the agents is broken down into a series of financially determined decisions affecting the structure of the AH as a commercial company (size and own factors), and another series of agro-economically determined decisions affecting the operation of the AH as an agricultural production system (allocation of own and external productive factors to specific agro-livestock activities).

The model allows the iterative simulation of several consecutive agricultural campaigns, allowing the dynamic evolution of the agents not only between the baseline year and the final year, but also throughout the transitional period of implementation of the public policy in question. The beneficiaries of the tool and model are eminently the policy-makers of European public administrations at supra-national, national and regional level, interested in carrying out impact analyses of implemented public policies and/or of the different policy options that are considered for their future implementation.

Potential users of the tool and model are the modellers of the technical institutions associated with the above public administrations, who advise policy-makers with their impact assessment reports.

References

[1] European Commission. Directorate General for Research and Innovation. and European Commission. Group of Chief Scientific Advisors., Scientific advice to European policy in a complex world. LU: Publications Office, 2019. doi: 10.2777/80320.

[2] S. Scheller, “Steven F. Railsback and Volker Grimm, Agent-Based and Individual-Based Modeling. A Practical Introduction,” oeconomia, no. 9–2, pp. 407–413, Jun. 2019, doi: 10.4000/oeconomia.5533.

[3] C. M. Macal and M. J. North, “Tutorial on agent-based modelling and simulation,” Journal of Simulation, vol. 4, no. 3, pp. 151–162, Sep. 2010, doi: 10.1057/jos.2010.3.

[4] K. H. Dam, I. Nikolic, and Z. Lukszo, Eds., Agent-Based Modelling of Socio-Technical Systems. Springer Netherlands, 2013. doi: 10.1007/978-94-007-4933-7.

[5] A. Downey, Think Complexity. 2012.

[6] A. Troisi, V. Wong, and M. A. Ratner, “An agent-based approach for modeling molecular self-organization,” Proc. Natl. Acad. Sci. U.S.A., vol. 102, no. 2, pp. 255–260, Dec. 2004, doi: 10.1073/pnas.0408308102.

[7] J. Testa, K. Mock, C. Taylor, H. Koyuk, J. Coyle, and R. Waggoner, “Agent-based modeling of the dynamics of mammal-eating killer whales and their prey,” Mar. Ecol. Prog. Ser., vol. 466, pp. 275–291, Oct. 2012, doi: 10.3354/meps09845.

[8] A. F. Griffin and C. Stanish, “An Agent-based Model of Prehistoric Settlement Patterns and Political Consolidation in the Lake Titicaca Basin of Peru and Bolivia,” Structure and Dynamics: eJournal of Anthropological and Related Sciences, vol. 2, no. 2, Oct. 2007, doi: 10.5070/sd922003290.

[9] K. H. van Dam, Z. Lukszo, L. Ferreira, and A. Sirikijpanichkul, “Planning the Location of Intermodal Freight Hubs: an Agent Based Approach,” in 2007 IEEE International Conference on Networking, Sensing and Control, 2007. doi: 10.1109/icnsc.2007.372774.

[10] Institute of Medicine. 2015. *Assessing the Use of Agent-Based Models for Tobacco Regulation*. Washington, DC: The National Academies Press. https://doi.org/10.17226/19018.

[11] A. Perello-Moragues and P. Noriega, “Using Agent-Based Simulation to Understand the Role of Values in Policy-Making,” in Springer Proceedings in Complexity, Springer International Publishing, 2020, pp. 355–369. doi: 10.1007/978-3-030-34127-5_35.

[12] H. Dawid and M. Neugart, “Agent-based Models for Economic Policy Design,” Eastern Econ J, vol. 37, no. 1, pp. 44–50, Dec. 2010, doi: 10.1057/eej.2010.43.

[13] E. J. L. Chappin, L. J. de Vries, J. C. Richstein, P. Bhagwat, K. Iychettira, and S. Khan, “Simulating climate and energy policy with agent-based modelling: The Energy Modelling Laboratory (EMLab),” Environmental Modelling & Software, vol. 96, pp. 421–431, Oct. 2017, doi: 10.1016/j.envsoft.2017.07.009.

[14] D. Kremmydas, I. N. Athanasiadis, and S. Rozakis, “A review of Agent Based Modeling for agricultural policy evaluation,” Agricultural Systems, vol. 164, pp. 95–106, Jul. 2018, doi: 10.1016/j.agsy.2018.03.010.

[15] R. L. Axtell, “Why Agents? On the varied motivations for agent computing in the social sciences,” 2000.

[16] F. C. Billari, T. Fent, A. Prskawetz, and J. Scheffran, "Agent-Based Computational Modelling: An Introduction," in Agent-Based Computational Modelling, Physica-Verlag, pp. 1–16. doi: 10.1007/3-7908-1721-x_1.

[17] V. Grimm et al., "The ODD Protocol for Describing Agent-Based and Other Simulation Models: A Second Update to Improve Clarity, Replication, and Structural Realism," JASSS, vol. 23, no. 2, 2020, doi: 10.18564/jasss.4259.

[18] B. Müller et al., "Describing human decisions in agent-based models – ODD + D, an extension of the ODD protocol," Environmental Modelling & Software, vol. 48, pp. 37–48, Oct. 2013, doi: 10.1016/j.envsoft.2013.06.003.

[19] Saberi Moghadam S, Samsami Khodadad F, Khazaeinezhad V. An Algorithmic Model of Decision Making in the Human Brain. Basic Clin Neurosci. 2019 Sep-Oct;10(5):443-449. doi: 10.32598/bcn.9.10.395. Epub 2019 Sep 1. PMID: 32284833; PMCID: PMC7149951.

[20] C. Li, "The DNDC Model," in Evaluation of Soil Organic Matter Models, Springer Berlin Heidelberg, 1996, pp. 263–267. doi: 10.1007/978-3-642-61094-3_20.

[21] N. Brisson et al., "An overview of the crop model stics," European Journal of Agronomy, vol. 18, no. 3–4, pp. 309–332, Jan. 2003, doi: 10.1016/s1161-0301(02)00110-7.

[22] A. de Wit et al., "25 years of the WOFOST cropping systems model," Agricultural Systems, vol. 168, pp. 154–167, Jan. 2019, doi: 10.1016/j.agsy.2018.06.018.

For preparing this report, the following deliverables have been taken into consideration:

Deliverable Number	**Deliverable Title**	**Lead beneficiary**	**Type**	**Dissemination Level**	**Due date**
D3.2	AI-based farmer's behavioural foundation	IDE	Report	Public	M37
D3.4	Biophysical models linking capabilities for the ABM	IAPAS	Report	Public	M37
D5.2	AGRICORE Land Market Module	AKD	Report	Public	M34
D5.4	AGRICORE Market Module	AKD	Report	Public	M34
D6.3	Biophysical model connection modules	IAPAS	Report	Public	M34

ANNEX: ODD+D Template

Following the recommendation of [18] the tabular form of the ODD+D is included as an annex. Müller et al. argue that "using the template makes the creation of an ODD+D description easier, since some categories can be answered by keywords such as "yes" or "no" instead of full sentences. The use of this tabular form simplifies the comparison of models applied in different studies to a large extent. In the main text, the overview and the design concepts should be copied and, if necessary, shortened".

Table 3: Description of AGRICORE Model using ODD+D template

Outline (à template)		**Guiding questions**	**Own ODD+D Model description**
I) Overview	I.i Purpose	I.i.a What is the purpose of the study?	The purpose of this model is to analyse at the microscopic level the effect of the implementation of the different monetary instruments and agri-environmental schemes associated with the Common Agricultural Policy of the European Union. In this sense, it allows both quantitative predictions to be obtained ex-ante to the implementation of agricultural measures and the corroboration of key performance indicators ex-post to their execution. AGRICORE, therefore, aims to become a fundamental tool for the design and evaluation of the impact of public agricultural policies.
		I.ii.b For whom is the model designed?	The model is fundamentally designed so that its results can be used to support the process of designing public policies by policymakers, both at the Community level in the EC and at the national level in the respective Ministries of Agriculture of the MS. However, its main users are likely to be modellers from the European Commission's own research centers (JRC) or scientists from private academic institutions dedicated to agroeconomics. They produce reports and studies based on the results obtained with the tool, which ultimately guide or support the decisions of policy-makers.
	I.ii Entities, state variables, and scales	I.ii.a What kinds of entities are in the model?	The model mainly considers individual farms (generically referred to as Agricultural Holdings), characterised by their constituent elements, namely: physical or legal personality, socio-demographic characteristics of the person(s) owning the farm, socio-demographic characteristics of the person managing the farm, amount of land available and allocation of activities on the farm, own mechanisation capacity, own labour capacity, and financial statement. Although the tool allows the generation of as many individual agents as there are real farms in a given area, for the analysis of certain measures, some types of farming may not be of interest to the modeller. However, they should still be considered in the simulation due to the effect of their interrelationships with the farms of interest under study. In these cases it is also possible to create super-agents (archetypical farms representing a set of farms of a certain typology or a complete production sub-sector). The markets for the exchange of productive factors (for the moment only land) and for the sale of output products from the AHs are also contemplated in the tool, not as agents but as external modules.

I) Overview			Additionally, the ecosystem of which the farms are a part can be represented by computing the effect on its biotic and abiotic factors of the actions of the agricultural holdings. However, so far, the ecosystem is not constituted as an agent in itself and its state (the mentioned effects) is not fed back into the AHs.
		I.ii.b By what attributes (i.e. state variables and parameters) are these entities characterized?	A detailed list of attributes is presented in Table 2. Decision-making by agents takes place on two clearly differentiated but interrelated conceptual-temporal planes (Figure 2). On the one hand, decisions affecting the holding structure itself (basically the size of the holding in terms of land operated or its capacity for self-supply of productive factors such as machinery or labour) are taken in the long term by assimilating the manager's strategy to a model-based predictive controller (MPC). On the other hand, in each simulation interval, assuming that the AH structure for that campaign is established, each agent makes decisions on the allocation of production factors to productive activities using an optimisation calibrated by means of possitive mathematical programing (PMP). Both steps are repeated iteratively for each simulation interval, after updating the agent's financial and agronomic states based on the yield and market realisation of the immediately preceding simulation step. The detailed description of these two optimisation-based behavioural models is presented in Deliverable D3.2.
		I.ii.c What are the exogenous factors / drivers of the model?	The main external factors in AGRICORE are: - the weather, which determines the performance of the different activities of the actors in each simulation step (campaign). - the policy instruments active during the simulation, which constrain the agent's space of possible actions and modify his preferences for some activities over others. - the interest rates that modify the agents' ability to afford and repay loans. - market prices for productive inputs and outputs.
		I.ii.d If applicable, how is space included in the model?	At the time of initialisation, each agent is assigned the code of the NUTS 3 to which it belongs, as well as a georeferenced location in Latitude and Longitude. This location is used to determine which other agents constitute the agent's neighbourhood and can be influenced by its actions, as well as in which land market each agent participates.
		I.ii.e What are the temporal and spatial resolutions and extents of the model?	Each simulation step represents one agricultural campaign. A priori, this could coincide with an agricultural year, but it could also be reduced in the case of use cases with more than one crop per year. In general, a simulation step occurs each time holding companies have to make new decisions on resource allocation to productive activities. In each simulation step, the productivities of each activity are also computed as a function of the given climatic conditions and the prices that each holding receives for its total production.

			Talking about spatial resolution in AGRICORE is complicated, because referring exclusively to the agents, the minimum distance between the locations of a pair of agents can be arbitrarily small. However, to calculate the annual yield, the biophysical models associated with the tool need both weather conditions and soil quality data. If we take into account that the spatial resolution of the models is never smaller than 25x25km we could consider this as the global spatial resolution of the whole tool.
	I.iii Process overview and scheduling	I.iii.a What entity does what, and in what order?	The complete sequence of the simulation flow using AGRICORE is shown in Figure 3-Figure 5.
II) Design Concepts	II.i Theoretical and Empirical Background	II.i.a Which general concepts, theories or hypotheses are underlying the model's design at the system level or at the level(s) of the submodel(s) (apart from the decision model)? What is the link to complexity and the purpose of the model?	The hypothesis of this model is that model-based predictive control (MPC) can be used as an algorithmic model of decision-making in the human brain [19]. This implies that, in general, two different models can co-exist: - A detailed and complex (generally non-linear) one, based on the economic-biophysical equations that determine the functioning of a farm as a commercial company and as an agricultural production unit. This model is the one used to simulate, season by season, the effect of the actions decided by the Farm Manager. - And another model, which can be less detailed and simpler (for example linearised around some break-even point), which is used for optimisation and which represents the (simplified) idea that the Farm Manager has of how the financial and agronomic dynamics of the agricultural holding works, i.e. the mental model that allows him to predict what will happen to the financial and agronomic states of the farm if he applies one or other management decisions. Therefore, the decisions made by agents at the structural and financial level are inherently bounded rational, due to both cognitive limitations (simplified mental model) and imperfect information (presence of unknown a priori and uncontrollable external shocks).
		II.i.b On what assumptions is/are the agents' decision model(s) based?	There are also some ad-hoc behavioural rules in the model, especially in relation to succession in the agricultural business. It is assumed that, at a certain pre-set age, farm owners retire. If they have descendants, they will continue the operation of the farm as long as the average annual income earned by the farm is higher than the average wage in the secondary and tertiary sectors in the geographical area where the farm is located.
		II.i.c Why is a/are certain decision model(s) chosen?	
		II.i.d If the model / a submodel (e.g. the decision model) is based on empirical data, where does the data come from?	As the AGRICORE tool intends to be an exhaustive microscopic tool able to assess the impact of very disparate policy instruments, there might be a lack of information to initialise some agent's attributes or modules' elements. When this happens an 'information gap' is detected. AGRICORE has proposed (deliverables D1.7 and D1.8) a methodology to fill these information gaps through different kinds of Participatory Research activities. In the particular case of the three Use Cases contemplated in the AGRICORE proposal, three respective survey campaigns have been implemented in Andalusia (Spain), Greece and Poland to disclose information leading to the

II) Design Concepts			quantification of parameters such as risk aversion, grade of innovativeness, technology transition costs, etc.
		II.i.e At which level of aggregation were the data available?	The AGRICORE tool has extensive data requirements, the level of aggregation of which is very heterogeneous: - The data used to construct the synthetic populations of agents come primarily from the FADN microdata, and are therefore a statistical sample at the level of individual agricultural holdings. However, some attributes are generated from socio-demographic data sources that may be at the level of municipalities (LAU2) or provinces (NUTS3). - Data from Participatory Research are at the level of statistical sample of individual farms. - Climate and soil data used by biophysical models and other impact assessment modules (IAMs) are usually aggregated at the level of a geo-defined grid, or in case of equivalent administrative unit resolution at NUTS3 or NUTS2 resolution.
	II.ii Individual Decision Making	II.ii.a What are the subjects and objects of decision-making? On which level of aggregation is decision-making modeled? Are multiple levels of decision making included?	Decision-making is modelled at two different levels, the political level and the agricultural level. The alternative policy instruments that can be implemented by policy-makers at different scales (European, national, regional), whose impact is to be analysed or predicted, constitute external disturbances for the agents. It is up to the latter to decide whether to adhere to these financial instruments (or agri-environmental schemes), provided that their status (set of financial-agronomic states) allows them to do so.
		II.ii.b What is the basic rationality behind agents' decision-making in the model? Do agents pursue an explicit objective or have other success criteria?	At the structural level (long-term financial optimisation), agents have two explicit objectives, which are to maximise net profitability (the ratio of profits to equity) while keeping the solvency ratio within safety margins. Fundamentally, this translates into determining the appropriate size of their farm (utilised agricultural area) and their own resources (machinery and/or permanent workforce). Each agent prioritises one objective or the other depending on its behaviour profile (i.e. depending on attributes such as risk aversion). At the agronomic level, the agents try to optimise the mix of agro-livestock activities in order to maximise the expected profit from the sale of the agricultural products generated during the season.
		II.ii.c How do agents make their decisions?	Agents take their decision at both levels by solving two different optimisation problems, posed by the respective objective functions and a set of constraints.
		II.ii.d Do the agents adapt their behavior to changing endogenous and exogenous state variables? And if yes, how?	Agents do adapt to changing exogenous variables and/or external shocks. For example within the decision of renting in/out. Based on the market prices, the agent will decide whether to rent or not. Likewise concerning buying/selling land. Another example is the modification of the decision-making due to agricultural policies or incentives (e.g. subsidies for ecological farming).

II) Design Concepts		II.ii.e Do social norms or cultural values play a role in the decision-making process?	At the moment, neither social norms nor cultural values are covered by the AGRICORE tool. However, agents belonging to the same geogrpahical vicinity may know the actions taken by their neighbours and (at least partially) the results of these actions, so that eventually terms could be added to the individual cost function in order to align the individual practice of each agent to the best practices observed at the local or regional level.
		II.ii.f Do spatial aspects play a role in the decision process?	Spatial location indirectly influences the decision making process of actors, as the quality of the soil that each actor owns (given by geo-referenced soil quality data) determines the average expected yield for different activities. Additionally, the simulated climatic conditions also vary according to the geographic location of the agents, indirectly influencing their future decisions.
		II.ii.g Do temporal aspects play a role in the decision process?	In this initial version of AGRICORE, agents have no memory (i.e. they do not take into account the degree of performance of past actions when making the next decisions). However, time does play a role in the optimisation process. On the one hand, as MPC is a sliding horizon technique, it requires a dynamic prediction model. On the other hand, in the objective function itself, and depending on the risk aversion profile of the agent, immediate profits or the expectation of higher future profits can be prioritised.
		II.ii.h To which extent and how is uncertainty included in the agents' decision rules?	The different level of risk aversion and the different levels of propensity to innovate are taken into account by the respective coefficients that weight the terms of the objective function, representing, for example, a different penalty for the same deviation of the actual solvency level from the target solvency level. Beyond that, the model does not currently include other mechanisms for incorporating uncertainty into the decision process. However, the very nature of the EMPC makes its stochastification relatively straightforward by including alternative future scenarios and weighted optimisation according to the probability level associated to each scenario.
	II.iii Learning	II.iii.a Is individual learning included in the decision process? How do individuals change their decision rules over time as consequence of their experience?	Learning has not been considered so far in the AGRICORE tool, partly because the initial period foreseen for the simulation of the use cases is $N_h = 7$years, which coincides with the usual period of validity of each Rural Development Plan. This timespan is too short for the agent to be able to draw the consequences of its (non) adherence to a given policy instrument and change its actions accordingly.
		II.iii.b Is collective learning implemented in the model?	For similar reasons, collective learning has not been considered in the tool for the time being either, beyond the possibility of agents imitating best practices at the local level.
	II.iv Individual Sensing	II.iv.a What endogenous and exogenous state variables are individuals assumed to sense and consider in their decisions? Is the sensing process erroneous?	Agents (Agricultural Holdings) perceive the average cost of land exchange from the Land Market Module (LMM) and use it for defining their next Land trading intentions. They also exogenously receive (or endogenously generate) output price prediction sequences.
		II.iv.b What state variables of which other individuals can an individual	Agents exchange information both directly and indirectly.

II) Design Concepts		perceive? Is the sensing process erroneous?	Through the Land Market Module, agents participating in the same local market house (i.e. agents belonging to the same municipality or geographical vicinity) can "hear" to the offers made by other agents for certain plots of land, thus being able to intuit what their unit land valuation is. Additionally, agents exchange technological information at regional level, by including the concept of "frontier farm" in the short-term model, as explained in deliverable D3.3.
		II.iv.c What is the spatial scale of sensing?	
		II.iv.d Are the mechanisms by which agents obtain information modeled explicitly, or are individuals simply assumed to know these variables?	The mechanisms for obtaining information from other agents have not been explicitly modelled, beyond the Land Market Module's own architecture, based on discrete auctions with public bids (which are known to all participants, who can adapt their future bids in response). However, no models of diffusion mechanisms for innovations or best practices have been explicitly incorporated so far.
		II.iv.e Are costs for cognition and costs for gathering information inclu-ded in the model?	Therefore, costs associated to information gathering or cognition are not considered either.
	II.v Individual Prediction	II.v.a Which data uses the agent to predict future conditions?	As agricultural managers do in real life, agents use exogenous forecasts from the relevant institutions to predict future prices and weather conditions.
		II.v.b What internal models are agents assumed to use to estimate future conditions or consequences of their decisions?	At the financial-structural level, agents rely on the MPC prediction model, which is nothing more than a simplified (possibly linearised) version of the non-linear dynamic model given by the equations in the #Dynamics section.
		II.v.c Might agents be erroneous in the prediction process, and how is it implemented?	Prediction errors are inherently covered by the use of the MPC itself. Deviations between the predicted state and the state actually achieved can be due to both differences between the prediction model and the simulation model (representing the limited cognitive capacity of the AH managers) and external disturbances applied during the simulation of each cropping season (representing the effect of uncertainties that the AH cannot accurately predict nor of course control).
	II.vi Interaction	II.vi.a Are interactions among agents and entities assumed as direct or indirect?	Both types of interactions are used in the tool. Direct interactions among agents occur in the Land Market Module while buying/selling land before each agricultural campaign. Indirect interactions occur in the exchange of technology through the use of the hypothetical frontier farm (D3.3).
		II.vi.b On what do the interactions depend?	The possibility for an agent to interact with other agents depends fundamentally on the geographical location of all of them, either because they belong to the same administrative area (municipality, agricultural district, etc.) or because they are all within a certain range of physical distance.

			In principle, these interactions are not limited by the typology of agents, and exchanges of land and technology are possible between any pair of agents.
		II.vi.c If the interactions involve communication, how are such communications represented?	The interactions that take place between agents within the Land Market Module are modelled as explicit bid or ask messages that are broadcasted to all agents belonging to the same local market.
		II.vi.d If a coordination network exists, how does it affect the agent behaviour? Is the structure of the network imposed or emergent?	There are no coordination networks between actors that modify their community behaviour.
	II.vii Collectives	II.vii.a Do the individuals form or belong to aggregations that affect, and are affected by, the individuals? Are these aggregations imposed by the modeller or do they emerge during the simulation?	Agents do not belong to or form any collectives.
II) Design Concepts		II.vii.b How are collectives represented?	
	II.viii Heterogeneity	II.viii.a Are the agents heterogeneous? If yes, which state variables and/or processes differ between the agents?	The heterogeneity of the actors is one of the main characteristics of the AGRICORE tool. During initialisation, in the process of generating the synthetic population, each agent is endowed with its own values for its attributes, while ensuring that the statistical distribution of the values of these attributes reproduces up to a certain minimum threshold of adjustment the statistical distribution of the values of these attributes in the real population. The main states that distinguish one agent from another are the economic dimension (given by the value of their gross output), their physical size (given by the amount of land they own), and the type of farming (given by the distribution of land to agricultural activities).
		II.viii.b Are the agents heterogeneous in their decision-making? If yes, which decision models or decision objects differ between the agents?	All actors have the same objects of decision: the development of their structure (own productive capacity) and the optimal allocation of these resources to specific agricultural activities. Although all agents share the same structure of optimisation problems that define their behaviour, some aspects of these problems may vary between agents: - The size of the agents (their economic size) may determine different financial capacity (e.g. with respect to the maximum degree of indebtedness, or with respect to the target solvency value). - A policy may be applicable only for a subset of agents that meet a set of requirements (whose states are between a certain range of values). In this case, the objective function or constraints defining the optimisation problem will be modified only for that subset of agents.
	II.ix Stochasticity	II.ix.a What processes (including initialization) are modeled by assuming they are random or partly random?	Subsequently, during the simulation, the weather conditions observed during each campaign can be set deterministically (this will be the case for ex-post impact analyses where historical weather observation data will already be available) or they can be generated randomly (this will

			be the case for ex-post impact analyses where no observations exist because of future campaigns).
	II.x Observation	II.x.a What data are collected from the ABM for testing, understanding, and analyzing it, and how and when are they collected?	Data on the actions taken by each agent (land acquisition and alienation, loan taking and repayment, selection of productive activities) and on the effects of these actions on their states are stored after each simulation interval, equivalent to a full agricultural season. This allows their partial visualisation during the simulation itself. However, the full usefulness of the collected data is not obtained until the simulation of the multi-year period is completed. It is then when, with all the data for each season, the analysis of the dynamic variation of the financial and agronomic statements at the individual and sector average level can be made. The aggregated data can also be used to calculate the socio-economic, environmental and ecosystem service impact through the corresponding Impact Assessment Modules.
		II.x.b What key results, outputs or characteristics of the model are emerging from the individuals? (Emergence)	The key results that emerge from the full use case simulation are the level of adherence of the agents to the policy measure(s) under analysis, and furthermore, the effect that these policies have on the financial and socio-economic states of each agent, with respect to the baseline scenario which would be the maintenance of the existing policies at t=0 and/or the implementation of no additional policy modifications. At the global level, the effect that the variation in agricultural activities produced by the tested policies has on the environment (variation in CO2 emissions, water use and pollution) and ecosystem services (variation in the area of habitats suitable for animal species, variation in animal populations, complementary activities to agriculture, etc.) emerges.
III) Details	II.i Implementation Details	III.i.a How has the model been implemented?	The main programming language used to implement the different modules and to exchange information among them is Python 3. R language is also used for the extraction of Bayesian Networks within the SPG process. In addition, for the mathematical modelling of the ST optimisation, GAMS is used as programming language and IDE. For solving the optimisation problems faced by the agents, external solvers such as Gurobi and MOSEK are interfaced.
		III.i.b Is the model accessible and if so where?	All the project deliverables and technical documents will be available at the project's website. The code is open-sourced at the project's GitLab public repository. Part of the generated data is available at the AGRICORE Zenodo open-data repository.
	III.ii Initialization	III.ii.a What is the initial state of the model world, i.e. at time t=0 of a simulation run?	At time t=0 of the simulation, the synthetic population of agents reproduces at a global level the real population on which the effect of the policies wants to be assessed. That is, the probability density distribution of the value of each attribute among the agents of the synthetic population reproduces (up to a certain level of adjustment) the probability density distribution of that attribute in the individuals of the real population.

III) Details			This is achieved by using Bayesian networks for the sequential generation of values and for the conditional generation of values for those attributes that correlate with each other (see deliverables D2.3 and D2.4)
		III.ii.b Is initialization always the same, or is it allowed to vary among simulations?	A priori, the composition and values of a synthetic population should not vary between two different simulations. This would not make sense as it would make it difficult to determine whether the different impact observed when simulating two different policies is due to the difference between the policies themselves or to the baseline difference between the initial populations of agents. What can vary are the weather conditions that will be used to simulate each season (and on which the yield and therefore the economic performance of the agents will depend). These conditions can be initialised deterministically or pseudo-randomly and can be different between two consecutive simulations of the same policy measure (e.g. to see the level of resilience of the agents adhering to the policy instrument under different climate scenarios).
		III.ii.c Are the initial values chosen arbitrarily or based on data?	
	III.iii Input Data	III.iii.a Does the model use input from external sources such as data files or other models to represent processes that change over time?	
	III.iv Submodels	III.iv.a What, in detail, are the submodels that represent the processes listed in 'Process overview and scheduling'?	
		III.iv.b What are the model parameters, their dimensions and reference values?	
		III.iv.c How were submodels designed or chosen, and how were they parameterized and then tested?	