https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

InsNet-CRAFTY v1.0: Integrating institutional network dynamics powered by large language models with land use change simulation

Yongchao Zeng[1], Calum Brown [1,2], Mohamed Byari[1☆], Joanna Raymond[1☆], Thomas Schmitt[1☆], Mark Rounsevell[1,3,4]

[1] Institute of Meteorology and Climate Research, Atmospheric Environmental Research (IMK-IFU), Karlsruhe Institute of Technology, 82467 Garmisch-Partenkirchen, Germany

[2] Highlands Rewilding Limited, The Old School House, Bunloit, Drumnadrochit IV63 6XG, UK

[3] Institute of Geography and Geo-ecology, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany

[4] School of Geosciences, University of Edinburgh, Drummond Street, Edinburgh EH8 9XP, UK

☆ These authors contributed equally to this work.

Correspondence: Yongchao Zeng (yongchao.zeng@kit.edu)

Abstract: Understanding and modelling environmental policy interventions can contribute to sustainable land use and management but is challenging because of the complex interactions among various decision-making actors. Key challenges include endowing modelled actors with autonomy, accurately representing their relational network structures, and managing the often-unstructured information exchange. Large language models (LLMs) offer new ways to address these challenges through the development of agents that are capable of mimicking reasoning, reflection, planning, and action. We present InsNet-CRAFTY (Institutional Network – Competition for Resources between Agent Functional Types) v1.0, a multi-LLM-agent model with a polycentric institutional framework coupled with an agent-based land system model. The numerical experiments simulate two competing policy priorities: increasing meat production versus expanding protected areas for nature conservation. The model includes a high-level policy-making institution, two lobbyist organisations, two operational institutions, and two advisory agents. Our findings indicate that while the high-level institution tends to avoid extreme budget imbalances and adopts incremental policy goals for the operational institutions, it leaves a budget deficit in one institution and a surplus in another unresolved. This is due to the competing influence of multiple stakeholders, which leads to the emergence of a path-dependent decision-making approach. Despite errors in information and behaviours by the LLM agents, the network maintains overall behavioural believability, demonstrating error tolerance. The results point to both the capabilities and challenges of using LLM agents to simulate policy decision-making processes of bounded rational human actors and complex institutional dynamics, such as LLM agents' high flexibility and autonomy, alongside the complicatedness of agent workflow design and reliability in coupling with existing programmed land use systems. These insights contribute to advancing land system modelling and the broader field of institutional analysis, providing new tools and methodologies for researchers and policy-makers.

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

1. Introduction

Scientists have developed various models to study land systems, given their critical role in exploring key topics such as climate mitigation pathways (Duffy et al., 2022), carbon storage (Ekholm et al., 2024), human fire use (Perkins et al., 2024), and land cover change (Calvin et al., 2022; Chen et al., 2019). Land systems encompass both natural and human factors, with policy interventions playing a pivotal role in shaping land use dynamics. These interventions serve as critical mechanisms for addressing climate change, preserving biodiversity, and ensuring food security (Broussard et al., 2023; Guo et al., 2024; Qi et al., 2018). The formation and implementation of land use policies are the product of complex institutional dynamics and can involve a wide range of actors with differing objectives and powers (Davidson et al., 2024), as well as multi-level governance systems, such as that of the European Union (EU) (González, 2016). Understanding how these actors interact and public policies evolve is crucial for assessing how changes in policy can influence the land system in the future, and what this can mean for environmental goals.

Despite the importance of being able to simulate the effects of institutional dynamics on land systems, and despite ample empirical evidence highlighting interconnectivity among institutional actors (Ariti et al., 2019; Díez-Echavarría et al., 2023; Tesfaye et al., 2024), there is a scarcity of land use models which incorporate institutional networks, due to the challenges of representing heterogeneous, autonomous institutional decision-makers. Among the few studies that have explicitly modelled institutional actors within the land system are González (2016) and Holzhauer et al. (2019). In these examples, institutional agents are rule-based and programmed to take limited actions in response to specific land use changes. To strengthen the connection between modelling and real-world policy-makers, Zeng et al. (2024b) developed an endogenous institutional model using a fuzzy logic controller mechanism that can integrate real-world policy-makers' knowledge as IF-THEN rules. Other studies employ the network of action situation (NAS) approach (Kimmich et al., 2023), which is developed from action situation and game theory (McGinnis, 2011), allowing for systematic integration of a wide range of empirical evidence. However, NAS is still in its infancy (Tan et al., 2023), and it does not yet offer an approach to create formalized models.

These studies have advantages in modelling specific aspects of policy institutions. However, we contend that advancing the holistic representation of institutional actors in formal models needs to overcome three key challenges: agent autonomy, complex relational structures, and unstructured data. Firstly, modelling institutional actors' autonomy requires accounting for heterogenous behaviour (Dakin and Ryder, 2020), involving learning and memory (Nair and Howlett, 2017) together with bounded rationality (Jones, 2003; Simon, 1972). Secondly, there are both horizontal and hierarchical structures in the policy-making process, which can result in complex relationships between institutional actors and a lack of clarity in the policy formulation process (Cairney et al., 2019). For example, within the EU, there are multiple scales and layers of governance and authority, existing alongside NGOs and lobbyists (González, 2016). Thirdly, modelling institutional networks is confounded by the unstructured nature of the data that are available to policy actors (Lawrence et al., 2014). Data can be textual, and come in diverse formats, such as policy documents, grey literature, and research papers, which require institutional actors to understand natural languages including technical language. These challenges are not unique to this field; the simulation of human behaviour or ecological dynamics in the land system is similarly

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

complicated, and solutions applied in these cases might be relevant here. Another similarity is in the value of such solutions, which cannot render a complicated system fully predictable but can reveal important dynamics stemming from behavioural processes (Davidson et al., 2024).

Large Language Models (LLMs), a form of artificial intelligence (AI), are based on numerous parameters that have been pre-trained on massive textual data and are designed to conduct natural language processes to understand and generate human-like text. The transformer architecture based on neural networks enables the LLMs to process sequences of text and contextual relationships between words (Vaswani et al., 2017). The text that LLMs produce is usually broken down into tokens, representing characters, sub-words or words (Minaee et al., 2024). LLMs have demonstrated strong language understanding and generation abilities and have emergent abilities such as multi-step reasoning that breaks down complex tasks into intermediate reasoning steps (Minaee et al., 2024). Hence, LLMs can be a powerful cognitive engine for autonomous agents that are able to sense the environment and act with regard to their own prescribed agenda (Wang et al., 2024). LLM agents' ability to process and understand natural language allows them to synthesize information from various sources including unstructured data.

LLM agents provide high flexibility in modelling complex interactions between multiple decision-makers. Park et al. (2023) simulated an artificial village with 25 villagers powered by LLMs. The simulated villagers had heterogeneous persona's and could interact with one another and their environment. These artificial villagers displayed believable, human-like behaviour and were able to organize a Valentine's Day party proposed by a user-controlled villager agent. Similarly, Qian et al. (2024) used LLM agents to simulate different roles in a software development team that is able to produce software cooperatively via a waterfall model. Further frameworks for dealing with many interacting agents have been emerging (see e.g., AutoGPT (Yang et al., 2023), AutoGen (Wu et al., 2023), AgentLite (Liu et al., 2024), MetaGPT (Hong et al., 2023)), which indicate the power of LLM agents in modelling complex relationships.

The aim of this study is to present a newly developed model InsNet-CRAFTY and explore the potential of modelling institutional networks in the land system using a state-of-the-art LLM agent approach. First, we identify the conceptual framework for implementing the institutional model and its coupling with a land use model. Specific tasks are assigned to the institutional agents to facilitate the interpretation and evaluation of the model. We analyse the agents' textual output and numerical output to evaluate the believability of their decisions and the resultant performance of their actions. We identify both opportunities and challenges for LLM agent applications in modelling institutional networks within the land system, which may provide useful insights into both model conceptualization and implementation for future research. This study also contributes to the broader field of institutional analysis in socio-ecological modelling, offering novel tools and methodologies for researchers and policy-makers.

2. Methodology

2.1 Model framework of InsNet-CRAFTY v1.0

We adopt the conceptual framework of a stylized, polycentric institutional network from González (2016), which offers a generic framework based on empirical evidence (e.g., peer-reviewed and grey literature) for Swedish forestry institutional actors. The key decision-makers included in the

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

conceptual framework are the government, research suppliers, environmental NGOs, (forest) owner associations, and supranational institutions. The government has three levels, namely national, regional, and local authorities. González (2016)'s framework features both hierarchical and horizontal structures, offering rich components of a polycentric institutional structure while maintaining parsimony for computational modelling.

We further adapt González's (2016) framework through generalisation and abstraction to obtain the conceptual framework for this analysis (see Fig. 1). The framework maintains González's (2016) structural features, but the hierarchical governments are abstracted into two layers with one comprising a high-level institution and the other several independent operational institutions (representing different policy sectors) leading to greater governmental polycentrism. Additionally, two new agents are included - a law consultant and a narrative injector. A description of all of the LLM agents follows here.

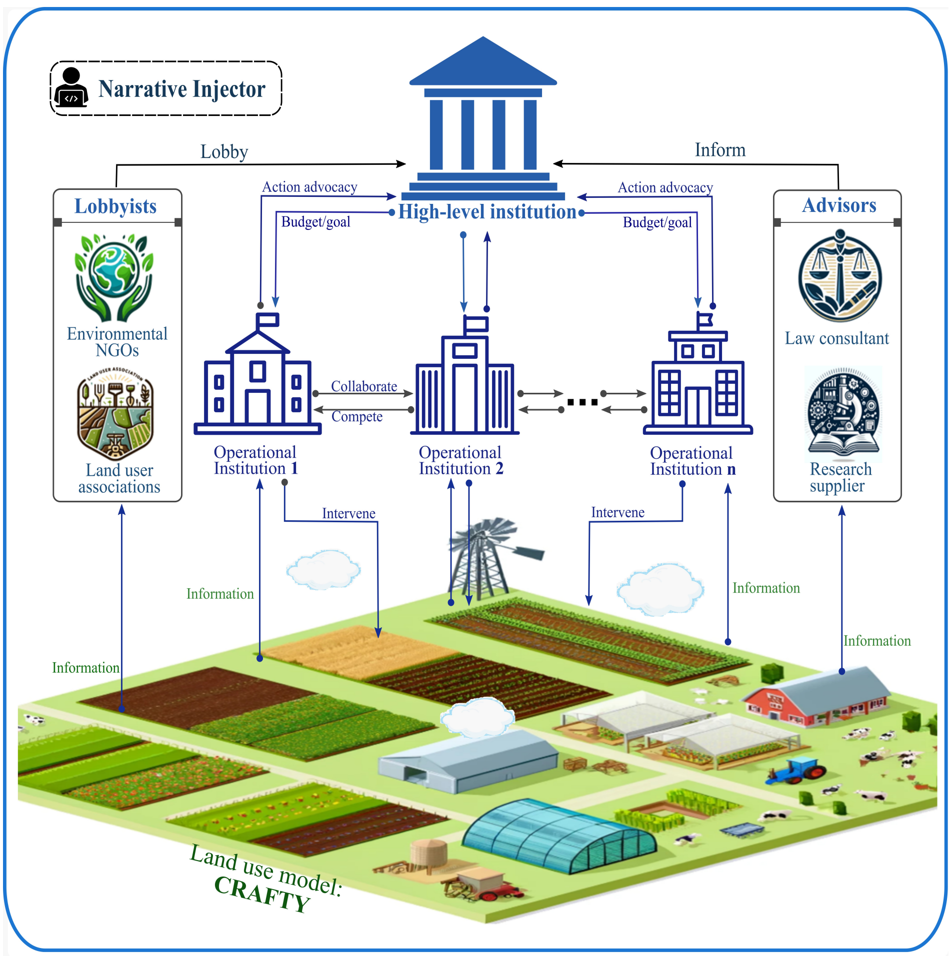

Figure 1: Conceptual framework for InsNet-CRAFTY v1.0. The institutional network model is adapted from Gonzalez et al. (2016) and coupled with the CRAFTY land use model (Brown et al., 2019). The hierarchical governments are abstracted into two layers with one comprising a

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

high-level institution and the other several independent operational institutions to achieve greater governmental polycentrism.

High-level institution: The high-level institution sets the overall policy ambitions and constraints (e.g. budgets) that affect the decisions of the operational institutions. The high-level institution aims to achieve mid to long-term policy goals based on the information provided by the operational institutions, research suppliers, lobbyists, law consultant, and narrative injector.

Operational institutions: Operational institutional agents represent different policy sectors (e.g., agriculture, nature conservation, forestry, transport). They adopt and execute concrete policy instruments to influence the decisions of land user agents in order to achieve specific policy goals. Operational institutions can also submit action advocacies to the high-level institution to obtain budgets or permissions for certain policy actions.

Lobbyists: Lobbyist agents represent professionals who advocate for specific interests or causes (e.g. environmental NGOs and land use associations). Lobbyists in the model can observe the state of the land use system and form their own opinions about what should be changed to reach their own objectives. Their advocacy is considered by the high-level institution when making policy adjustments.

Advisors: Advisory agents can inform the high-level institution's policy-making using professional knowledge and skills. The framework considers two types of advisors: research suppliers and law consultants. The research suppliers observe land use changes and provide a description of the current and possible future land use states. They analyse and interpret both numerical and textual data to support the high-level institution's decision-making. Law consultants offer information about existing laws, regulations, policies, etc., that legally underpin the high-level institution's policy actions; here we use EU policy documents to define these.

Narrative Injector (optional): An actor whose absence does not affect the functioning of an institutional network but can introduce highly unstructured exogenous disruptions into the model simulations through narratives (e.g., protest, war, unexpected disasters). The narratives can interact with all actors in the model and can be injected at any point during the simulation. The narrative injector provides the means to explore the impact of shock and extreme events on the functioning of the institutional model.

Together with these institutional agents, we apply the CRAFTY land use model (Brown et al., 2019; Murray-Rust et al., 2014) to simulate land use changes in response to the institutional agents' interventions and potentially other drivers of change, e.g. socio-economic and climate change. The LLM agents form a stylized polycentric institutional model that can be implemented in a sequential order. For instance, CRAFTY can produce information indicating that both meat supply and protected areas (PAs) need to be improved to achieve better food security and nature conservation. Then, the research supplier, operational institutions, and lobbyists collect and analyse the relevant data generated from CRAFTY. The data analysis serves as a basis for these agents to form different narratives that fit their distinct roles. The law consultant references policy and law documents to extract relevant information. The narrative injector may output a piece of news about an emergent incident. All these agents' output is eventually fed to the high-level institution, which considers

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

the different stakeholders' positions and strives to make balanced decisions. The high-level institution has concrete actions to influence the behaviour of the operational institutions, such as budget allocations and policy goal adjustments. The operational institutions utilize their budgets to formulate policy instruments such as subsidies, taxes, and administrative measures to steer meat supply and PA coverage towards the target levels. It is worth noting that the high-level institution does not have to be activated at the same frequency (in time) as the operational institutions, reflecting the asynchronous nature of agent decision-making at different levels. Appendix A provides extra details and a technical description of the model's sequential processes.

2.2 LLM agent framework

To implement the institutional network model, the agents have to be equipped with a powerful "brain". Because of the extremely rapid evolution in the LLM field, a variety of ways to create LLM agent "brains" have been emerging (Sumers et al., 2024). Here we use the framework in Fig. 2 to represent the cognitive architecture of an LLM agent, which derives from the unified framework proposed by (Wang et al., 2024), and the LangChain framework (LangChain, 2024).

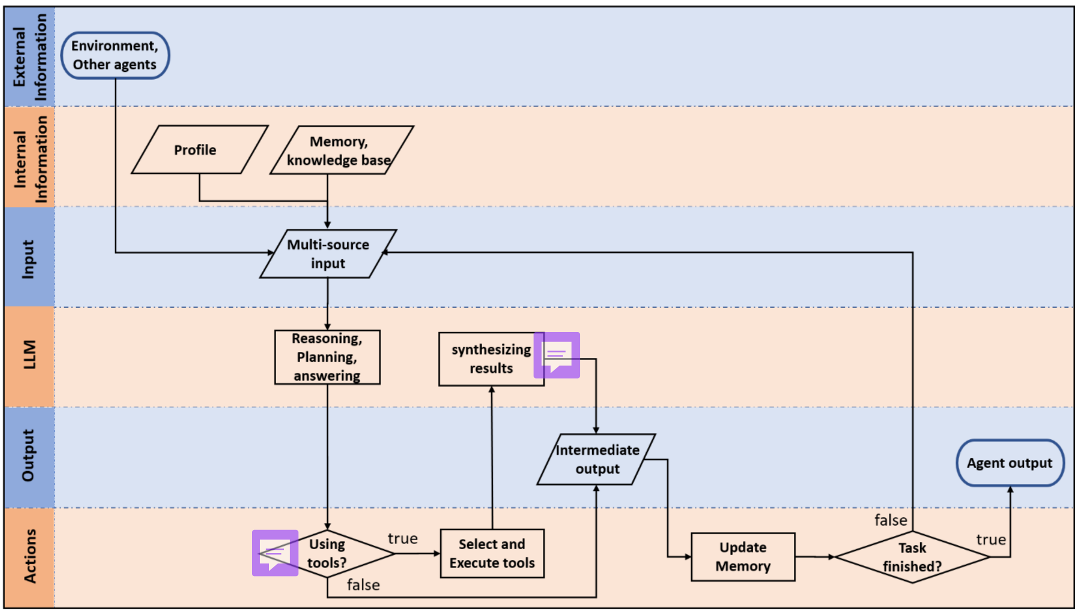

Figure 2: The cognitive architecture of an LLM agent. The core procedures of a LLM agent include the input, output, and the LLM. The agent's capability can be enhanced by integrating sophisticated work-flows such as memory, tool use, and reflection.

Although the complicatedness of different agents' "brains" varies, the core of a LLM agent consists of a LLM and the LLM's input and output. The functionality of the LLM agent can be enriched by incorporating more information into the input. Besides receiving external information from the modelled environment and other agents' responses, the LLM agent integrates internal

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

information such as a profile describing its identity, objectives, decision guidance, etc. An agent can also incorporate memory and a knowledge base into its input. Memory is divided into short-term memory and long-term memory. Short-term memory with high temporal relevance is embedded directly into the agent's prompt (the input directly received by an LLM). Knowledge and long-term memory relevant to a given decision-making context are extracted using Retrieval Augmented Generation (RAG) (Fan et al., 2024). This multi-source information forms the input to prompt the LLM to generate reasoning and planning, or to answer specific requests. If the agent is given a task to complete, the LLM helps to divide the task into small and achievable sub-tasks.

The capabilities of an LLM agent extend beyond text generation; it can actively execute sub-tasks and make decisions about the necessity of tools for task completion. In this context, "tools" refers to functions coded in programming languages such as Python. For instance, a function might perform calculations that current LLMs struggle to handle reliably on their own. An agent selects and employs appropriate tools to advance a task, as required. These tools process and organize results, which the LLM then synthesizes and outputs in natural language. Initially, these outputs are considered intermediate. The agent updates its memory by organizing and storing relevant inputs and outputs as necessary. Subsequently, it evaluates whether the tasks are complete to decide whether to produce the final output or to continue processing with updated memory.

2.3 Experimental Settings

The CRAFTY land use model is a crucial component of the simulation environment, within which the institutional agents operate. We set up the land use model according to CRAFTY-EU (Brown et al., 2019) and parametrized it with the data for the RCP2.6-SSP1 climatic and socio-economic scenario (Brown et al., 2019). The CRAFTY-EU model uses a map of European countries at a 5-arcminute resolution. The scenario simulation covers the period from 2016 to 2076. The data are available on Zenodo (Zeng et al., 2024c)

To enhance the focus of the experiments and facilitate the analysis, we narrowed the scope of the modelled actors by specifying their roles and responsibilities. Instead of integrating a diverse array of operational institutions with a wide range of policy objectives and tools, we incorporated two distinct operational institutions focused on different policy sectors: an environmental institution and an agricultural institution. The former prioritises environmental protection with a specific aim of expanding protected areas (PAs) for nature conservation, while the latter focuses on meat production to ensure food security using economic policy instruments such as subsidies. Since meat consumption is a major driver of deforestation, greenhouse gas emission, climate change, and biodiversity loss (Djekic, 2015; Machovina et al., 2015), and its consumption continues to increase (Petrovic et al., 2015), this experimental setting creates a conflicting context for the two institutions. They compete for limited budgets to fulfil their respective policy objectives.

Lobbyists actively seek to influence the high-level institution by advocating for increased financial support to either enhance the PAs or develop the meat industry. The research supplier analyses and interprets the data generated by the CRAFTY model; while the law consultant uses RAG to retrieve relevant information from a selected set of EU policies. The data are available on Zenodo (Zeng et al., 2024c). Consequently, the high-level institution is tasked with managing the interplay and potential conflicts between these agents, striving to balance budget allocations with the practical achievement of policy goals. The specific experimental purpose is to explore how these

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

agents reason and make plans in favour of their positions and to evaluate their performance in policy goal adjustments/achievements and budget allocation.

To improve the performance of the simulations, the lobbyists are allowed to use the output from the research supplier to strengthen their arguments. We chose not to incorporate the narrative injector agent in the results reported here for simplicity and in order to maintain the system's full autonomy. We followed an AI-assisted prompt development procedure depicted in Zeng et al. (2024a) and sought to use straightforward language to form the prompt templates. The prompt templates are given in Table B1 – B7.

As previously stated, the high-level institution and the operational institution are not synchronous. Here, the high-level institution is activated every ten iterations, while the operational institutions adjust their policies every two iterations, representing a more frequent response in policy adjustments. This frequent adjustment reflects the agility of the operational institutions compared to the slower, more deliberative pace of the high-level institution.

We set the initial target meat supply as 1.2 times the initial meat production level, and the target of PA coverage as 10% of the total land area. These parameters give the institutional actors slightly higher but achievable initial targets to pursue. The initial budget allocation is equally divided between the operational institutions.

To implement the model, we used a combination of different LLMs to power the agents to improve the token cost. The LLMs agents with actions were built using the LangChain library. The agents' features are summarized in Table 1. The equations that describe the high-level and operational institutions' non-LLM behaviour as well as related numerical settings can be found in Appendix C. The code is available on Zenodo (Zeng et al., 2024d)

Table 1: The experimental settings of the LLM agents

Agent	Input	Action	Output	LLM	Remarks
Law consultant	1)Document containing EU laws, policies, regulations etc. 2) Profile	Using RAG.	Unstructured text to inform the high-level institution's decision-making	Llama-3-70b-8192	Goals: Extracting relevant information from a knowledge base to inform the high-level institution's legal actions.
Research supplier	1) CSV file containing data from CRAFTY 2) Profile	Wiring and executing Python code to analyse the data.	1) Unstructured text to inform other agents 2) Intermediate output	gpt-4o	Goal: Analysing and interpreting the data generated by CRAFTY.

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

Environmental NGO	1) Research supplier's output 2) Profile	None	Unstructured text to lobby the high-level institution	Llama-3-70b-8192	Goal: Lobbying the high-level institution to prioritise nature conservation.
Land user associations	1) Research supplier's output 2) Profile	None	Unstructured text to lobby the high-level institution	Llama-3-70b-8192	Goal: Lobbying the high-level institution to prioritise meat industry development.
Environment-al institution	1) CSV file containing data from CRAFTY 2) Profile	Wiring and executing Python code to analyse the data.	1) Unstructured text to inform the high-level institution 2) Data from CRAFTY code	Llama-3-70b-8192	Goal: Striving to acquire budget to support PA expansions to reach the target PA coverages. Policy instrument: PA designation.
Agricultural institution	1) CSV file containing data from CRAFTY 2) Profile	Wiring and executing Python code to analyse the data.	1) Unstructured text to inform the high-level institution 2) Data from CRAFTY code	Llama-3-70b-8192	Goal: Striving to acquire budget to support meat production to reach the target meat supply level. Policy instrument: economic measures (e.g., taxes and subsidies)
High-level institution	1)Unstructur-ed text from all other agents 2) Profile	None	1) Unstructured text 2) Policy goal and budget allocation adjustments in JSON structure	gpt-4o	Goal: Making policy adjustments based on multiple stakeholders. Policy Instrument: Administrative orders to adjust the operational institutions' policy goals; financial measures to allocate budget between the operational institutions.

3. Results

3.1 Textual output

The LLM agents' output contained 19808 words (28778 tokens) and 48 plots. We summarised the textual output that demonstrates the behavioural patterns of the agents, while also highlighting counter-intuitive or potentially erroneous agent behaviours. This allows the agents' general behavioural regularities and occasional malfunctions to be displayed simultaneously.

3.1.1 The advisors' output

The research supplier completed various tasks, including checking missing values in the data, interpreting the trend of meat supply and demand, analysing the discrepancy between policy goals

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

and actual outcomes, budget allocation, and calculating the correlation between different time series. Figure 3 briefly illustrates the work-flow of the research supplier agent. It made plans, executed the plans step by step and interacted with tools. The tools offered returned values to form intermediate output, which was fed back to the LLM. A final output was produced based on the intermediate outputs. In some of the final outputs, the agent attached a note at the end of the output as a reminder of the applicable scope of the analysis, e.g., "Note: The above insights are based on the analysis of the provided data and may not be generalizable to other contexts". In 2066, the agent encountered an error -- "Agent stopped due to iteration limit or time limit" – indicating the actions the agent needed to take exceeded the set maximum execution time.

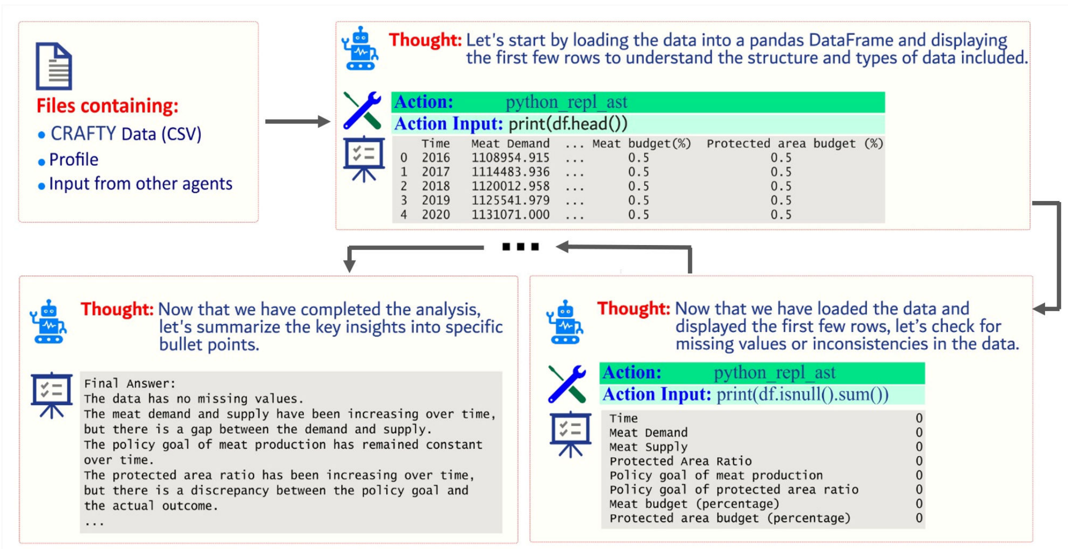

Figure 3: The work-flow of the research supplier agent. The agent took the initial input to generate a thought to decide what actions should be taken to analyse the data. Then, it executed the action by calling a function, which in turn produced the intermediate results. These results served as a part of the updated input to let the agent generate a new thought for the next iteration. After several iterations of thought-action-output loops, the research supplier agent produced a final interpretation of the data.

The law consultant emphasized six critical aspects to influence the decision-making of the high-level institution, based on the available knowledge base. These aspects include "biodiversity and ecosystem restoration targets," "agricultural production and environmental impact," and "climate change mitigation." The agent not only highlighted these issues but also cited relevant laws, policies, and directives. Furthermore, the agent elaborated on the implications of these legal and policy frameworks for the high-level institution's policy-making processes. For example, in discussing "biodiversity and ecosystem restoration targets," the law consultant noted that "the EU Restoration Law mandates the restoration of at least 20% of the Union's terrestrial and marine areas by 2030, and all ecosystems in need of restoration by 2050." This law was interpreted to mean that "a significant portion of the budget should be dedicated to protected areas to meet these objectives." It was observed that the law consultant agent produced the same output repeatedly

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

over several iterations, reflecting stagnation due to the absence of new inputs that could prompt different responses.

3.1.2 The lobbyists' output

The environmental NGO generated a variety of arguments for prioritising protected area establishment over meat production. For instance, in some years the agent highlighted the urgent need for nature conservation, the impact of meat production, or the necessity of budget increase. In 2066, the environmental NGO agent did not receive information from the research supplier due to the error mentioned above. However, this error did not paralyze the simulation. Instead, the LLM agent stated "I apologize, but it seems like there is no information provided. However, as a representative of an environmental NGO, I can still provide some general bullet points to lobby a high-level public policy institution to prioritise nature conservation". Without basing its arguments on data, the agent emphasized the economic benefits of nature conservation, the importance of PAs to climate change mitigation and adaptation, human health and well-being.

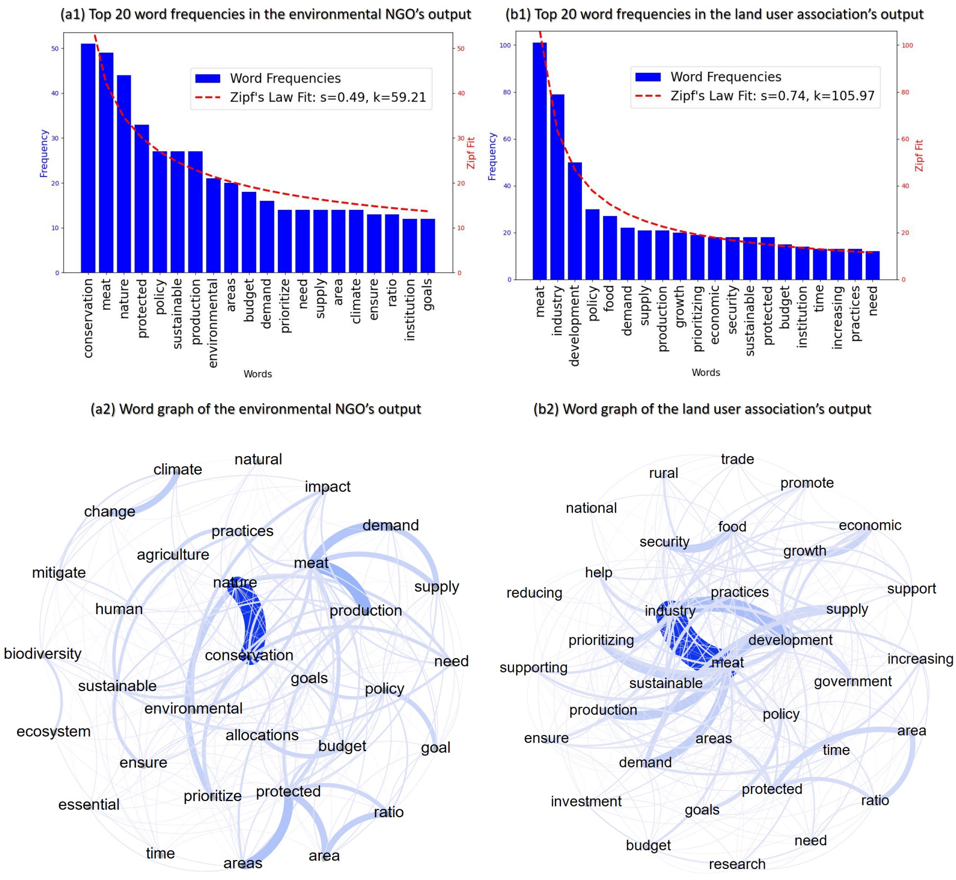

Figure 4: Word frequencies and word graphs derived from the lobbyists' output. The dashed red lines in (a1) and (b1) are derived by fitting Zipf's law distribution to the word frequency distributions. Zipf's law can be expressed as $f(r) = k/r^s$, where f(r) is the frequency of a word;

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

r represents the rank of the word according to its frequency; k and s are parameters. A larger s indicates a set of words distributed more unevenly. It can be seen in (a1) $s = 0.49$ for the environmental NGO's output and in (b1) $s = 0.74$ for the land user association's output, reflecting the two agents' different approaches in formulating their arguments. The word graphs only display nodes that have more than thirty links, in order to maintain visual clarity.

The land user association agent also utilised background information and the data interpretation provided by the research supplier agent to lobby the high-level institution to prioritise meat industry development. For instance, this agent highlighted economic growth, job creation, food security, and alignment with policy goals. When the output from the research supplier agent was missing, it gave more general bullet points to lobby the high-level public policy institution, including emphasizing the meat industry's economic benefits, food security, rural development, innovation and technology without using any data from CRAFTY.

The lobbyists had high autonomy to defend their interests but were not given detailed instructions about how to persuade the high-level institution. To better visualize how the lobbyists formulate their arguments, Fig. 4 illustrates the word frequencies and relationships through word graphs derived from their outputs. The analysis reveals a less prominent skew in the frequency distribution of the top 20 words used by the environmental NGO compared to those of the land user association. This can be quantified by the parameters of Zipf's law distributions fitted to the word frequency data. The environmental NGO frequently emphasized the term "conservation" and notably the word "meat." Its discourse primarily focused on two aspects: the environmental threats posed by meat production and the critical importance of conservation efforts. This concern was underscored by the research supplier's data interpretation showing a widening gap between meat demand and supply. In contrast, the land user association highlighted the development of the meat industry and food security, without opposing the expansion of protected areas. Instead, the land user association consistently advocated for sustainable meat production practices, which they argued would support their request for an increased budget.

3.1.3 The operational institutions' output

The agricultural institution's outputs consistently addressed the discrepancies between the meat production policy goals and the actual outputs, alongside recurring budget challenges. This agent repeatedly emphasized the necessity of addressing budget deficits, advocating for more efficient budget allocations and increased financial support to meet production goals. Key recommendations included increasing budget allocations to bridge the gap between policy goals and actual outcomes, setting realistic policy goals that align with current capacities, and enhancing sector productivity through various initiatives, e.g., farmer incentives and sustainable practices. Additionally, the institution suggested establishing a robust monitoring and evaluation framework to regularly assess the effectiveness of policies and adjust as necessary. A holistic approach was advocated to balance increased production goals with budget constraints, thereby boosting food security, improving farmer livelihoods, and ensuring financial well-being.

The environmental institution consistently highlighted a gap between the current state of protected areas and policy goals over the decades, emphasizing the need for increased financial support and a higher priority for protected area establishment to achieve biodiversity conservation and pollution reduction. Recommendations include raising the PA goals incrementally each year, improving governance, enhancing community engagement, and specifically allocating a substantial percentage of budget surpluses to facilitate the expansion of PAs. These steps were deemed crucial by this agent for reaching Net-zero targets and effectively managing biodiversity conservation amidst evolving environmental challenges. However, the agent mistakenly used mean values to describe the time series, which generated misleading outcomes. For instance, in the year 2076, the actual protected area is 25.14% and the target is 30.17%; however, the environmental institution used the mean values of 13.44% and 17.40% respectively to inform the high-level institution about the current situation. This error did not, however, qualitatively change the need to expand protected areas.

3.1.4 The high-level institution's output

The high-level institution employed a systematic and analytical approach to decision-making, consistently integrating stakeholder feedback across several sectors to refine policy goals and allocate budgets effectively. This process involves a detailed analysis of input from agricultural and environmental institutions, NGOs, and industry associations. Key actions include adjusting policy goals and redistributing budget percentages to better support the targeted outcomes in meat production and environmental protection. The institution regularly adjusted its strategies, intending to bridge the gaps between current outcomes and policy objectives, focusing on sustainability, economic stability, and nature conservation. However, the output of the high-level institution was sometimes inaccurate. For instance, the high-level institution's analysis only included information from all six of the LLM agents in 2036 and 2056 with the law consultant and/or the research supplier's inputs occasionally being missed.

3.2 Policy actions and outcomes

The results shown in Fig. 5 (a) illustrate that the high-level institution increased the policy goals of the PA coverage gradually across the simulated time period, which resulted in a stepped pattern of PA growth. This reflects the periodic activation of the high-level institution as described previously. The actual PA coverage seemed to be well controlled by the operational institution because the actual PA coverage shows a prominent tendency to approach the target PA coverage. The eventual policy goal was set at approximately 30%, which drove the actual PA coverage to approach this level. In some years (between 2046 and 2076), the actual PA coverage reached the target and then remained almost unchanged for several years until the high-level institution raised the targets again.

The variation in the gap between the target PA coverage and the actual coverage illustrates the environmental institution's tendency to follow the policy target. In the beginning, the gap between the target PA coverage and the actual PA coverage was large, but the gap shrank over time until in 2052 the gap diminished to approximately zero. Then the actual PA coverage stayed almost unchanged from 2052 to 2055, which indicates that the operational institution imposed negligible influence on the land use system to maintain a small target-outcome gap. In 2056, the high-level institution raised the policy target to form a notable gap again. The environmental operational institution continued expanding the PAs to minimize this gap. This pattern repeated and resulted in the stepwise shape of the time series of the gap. These results demonstrate the alignment of the high-level institution's policy goal adjustments with the environmental institution's capability to influence PA coverage. However, the budget surplus remained positive and grew over time, which indicated over-funding by the high-level institution.

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

Similar to the policy goal adjustments in the PA coverage, Fig. 5(b) shows that the high-level institution increased the target level of meat production periodically and gradually, resulting in a stepped growth over time. Meat supply is positively correlated with the policy goal. Although the meat supply was not able to reach the policy goal, the goal-supply gap was limited within a relatively small range. In 2065 meat supply plateaued, while the ensuing policy goal adjustment at 2075 was still increasing. In contrast with the environmental institution, the agricultural institution underwent increasingly severe budget restrictions.

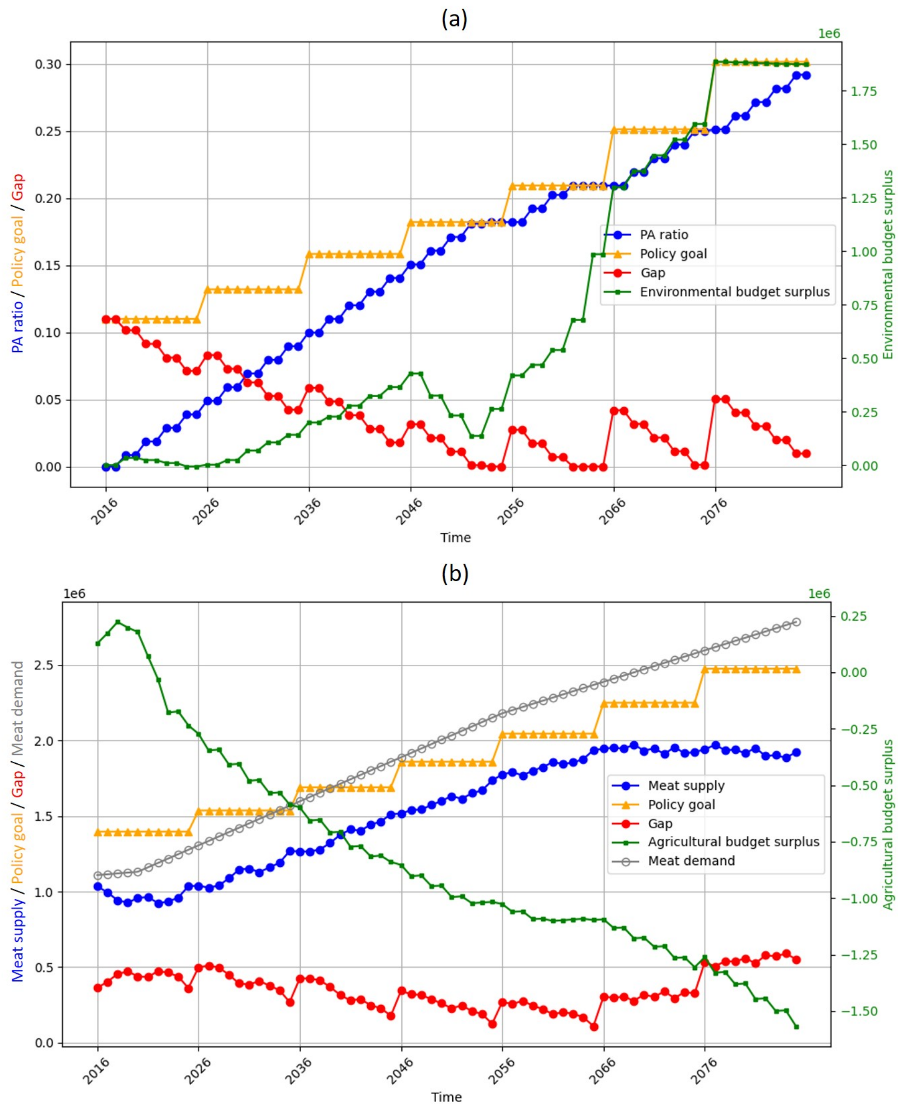

Figure 5: Policy goal adjustments, budget allocation, and their impacts for a) the environmental operational institution agent and b) the agricultural operational institution agent.

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

Figure 6 shows the budget allocated by the high-level institution. In the first ten years (from 2016 to 2025), the budget allocation between the two operational institutions is 50/50 by default. However, it can be seen that the high-level institution shows a tendency to avoid imbalanced budget allocation. Despite the agricultural institution's lack of budget, the budget allocated between these two operational institutions was 60/40 from 2026 to 2045, 30/40 from 2046 to 2055, and 45/55 from 2076 to the end of the simulation. The high-level institution chose to allocate more budget to the agricultural institution in only twenty iterations even if the latter's budget deficit occurred very early (before 2026).

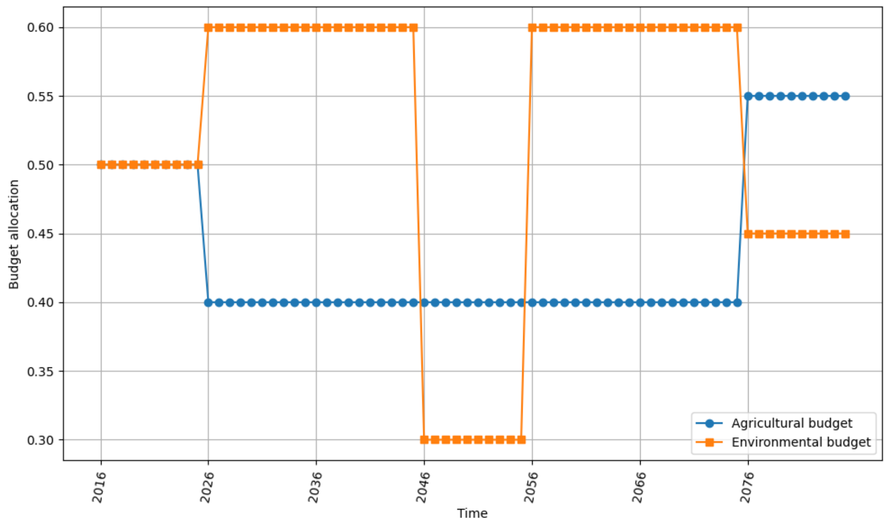

Figure 6: Budget allocation between the agricultural and environmental operational institution agents.

The twenty iterations include an unexpected budget drop for the environmental institution in 2046, leading the sum of the budget allocation ratio to be 0.7, when it should be 1.0 as in the other iterations. This sudden drop in budget is caused by the high-level institution's decision to allocate 30% of the total budget to "other initiatives and programs". This should result from the research supplier stating that "There are correlations between meat demand, meat supply, and protected area ratio, but it is not clear what the causative factors are". The research supplier's statement prompted the land user association to propose "further research is needed to understand the causative factors. We propose funding for research initiatives to better understand these relationships and inform evidence-based policy decisions". Although this unexpected drop in the budget allocated to the environmental institution demonstrates the coherence of information transmitted among multiple agents, it could also become an intractable issue for LLMs embedded within existing hard-coded systems.

4. Discussion

4.1 Believable behaviour of the LLM agents

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

Building upon the pre-trained LLMs, the institutional agents modelled in InsNet-CRAFTY exhibited diverse human-like reasoning and actions. The agents' behaviour was guided using prompts in natural languages, which gave the model developers high flexibility in creating the agents' autonomous behaviours. This flexibility facilitated the modelling of the complex relational structures among the agents. Given appropriate profiles, the agents were clear about their identities and relationships with others, demonstrating consistent role-specific decision-making. The LLM agents also showed an ability to handle the qualitative and unstructured information generated by the lobbyists, law consultant, and advocacies from the operational institutions. The capability of function calling (e.g. writing and executing computer code) further improved the agents' autonomy, enabling the latter to deal with more complex tasks such as data analysis and knowledge retrieval. These capabilities suggest that LLM agents have a unique potential to overcome the key challenges in modelling institutional networks.

Besides these apparent strengths, the LLM agents showed conflicting but understandable behaviour when faced with key real-world challenges such as conflicting objectives, uncertainty and imperfect (or absent) information. The budget allocation was a major output of the high-level institution, which reflected competing claims for a limited resource, as well as an impromptu suggestion by one of the lobbyist agents that money be transferred to research to better understand policy impacts. These dynamics could allow many important policy processes to be investigated, including observed differences between budgeting systems based on plurality, proportional representation or public participation, in which information inputs and decision-making powers vary substantially (Feindt, 2010; Hallerberg and Von Hagen, 1997; Lee et al., 2022).

The numerical results show that the environmental institution in our simulations was over-funded, while the agricultural institutions were struggling with an inadequate budget. We found that this imbalance was prompted by the environmental institution and the research supplier misleadingly informing the high-level institution that PA coverage was positively correlated with budget surplus. Indeed, both the two operational institutions insisted that their respective policy targets (PA coverage and meat production) should be increased because those targets were positively correlated with other desired outcomes. However, advocacy efforts were not equally effective. The environmental NGO's arguments were backed up by urgent environmental concerns, and outweighed the more economically focused arguments of the land user association. This imbalance might derive from the LLMs' training data being influenced by present-day social norms and highlights the potential for biases to be embedded within the agents' roles. These also of course reflect policy biases in reality, where norms, power relations, communication and urgency all affect policy priorities, and potentially allow exploration of approaches to mitigate these issues in differing policy contexts (Barnett and Finnemore, 1999; Sinden, 2004; Yami et al., 2019).

It can be hypothesised that the superficial use of correlation methods to interpret the data, the misleading arguments formed by the operational institutions, and the competition between the lobbyists all contributed to the high-level institution's path-dependent decision-making in both policy target adjustments and budget allocations. The lack of decisive action to fix the issue also implies that it is very challenging for the high-level institution to find an optimal solution given the need to consider many stakeholders' conflicting interests in the policy-making process. Nevertheless, such limitations do not necessarily diminish the LLM agent's usefulness in simulating human decision-making, rather it captures important and believable behaviour in terms of bounded rationality (Simon, 1972; Gigerenzer and Goldstein, 1996; Jones, 2003) and policy-makers muddling through (Lindblom, 1989) within complex systems.

4.2 Challenges of implementing LLM agents

Along with the advantages of the LLM agent approach in simulating institutional networks, erroneous behaviour was also apparent. Typical causes of errors were flaws in agent work-flows and LLM hallucinations (Ji et al., 2023; Perković et al., 2024; Yao et al., 2023). The research supplier agent's occasional failure to output data interpretation was caused by a flaw in the agent work-flow that generated the error "Agent stopped due to iteration limit or time limit.". This error could easily be avoided by increasing the permitted number of iterations that an agent needs to execute a complex task, although it had the advantage of forcing the other agents to proceed with imperfect (out-of-date) information, as is often the case in real-world contexts.

Unlike the research supplier, the operational institutions were not given specific data analysis instructions. This led to an unexpected outcome -- the operational institutions tended to use mean values to describe the latest state of the time series and so provided misleading information to the high-level institution. Such erroneous behaviour can be categorized as hallucination because the agent used plausible-sounding words to express nonsensical information (Ji et al., 2023). This erroneous behaviour could be mitigated by using more specified instructions in the prompts to guide agents' reasoning or designing extra mechanisms to detect and rectify the LLM's response (Tonmoy et al., 2024). However, addressing LLM hallucination is challenging, and there is no standard solution so far.

For large-scale, land use models, another crucial challenge is an LLM's limited context window. Here, the high-level institution had to consider all the other agents' output to make decisions. The resultant input could be very lengthy. Issues might arise if the input exceeds the maximum number of tokens (namely the size of the context window) that an LLM could digest. In the real world, institutional networks are far more complex than those in this model, and it is not unusual for high-level institutions to be overwhelmed by the information that they need to assimilate, or to use information selectively as a result (Bainbridge et al., 2022; Fischer et al., 2008; Rich, 1975). The limited size of the context window can therefore be seen as a feature that reflects human decision-makers' bounded rationality and information processing capabilities or preferences, as well as the imperfect nature of much information used for decision-making (Neri and Ropele, 2012). However, whether it is preferable to model such a feature in a controlled manner or to rely on the result of a technical limitation is debatable. The technical limitation could be mitigated by using summarized input or including memory mechanisms with retrieval methods (Modarressi et al., 2023; Zhong et al., 2024; Zhou et al., 2023), although these approaches all require extra effort in designing peripheral agent work-flows.

In contrast to the above errors, the data formatting issue could be more cumbersome to handle. Because the LLM agents were coupled with a programmed land use model, the LLM agents needed the capability to structure data in a designated format, otherwise, the programmed model would not parse the data, which could disrupt the simulations. Here, we used JSON, on which many current LLMs have been fine-tuned, to format the output of the high-level institution. However, there is no guarantee that LLMs can always accurately format their output. This leaves an extra task to design peripheral work-flows to secure the format. In this model, we employed three layers

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

of mechanisms to derive the correctly formatted data, including re-prompting the LLM, using regular expressions (Li et al., 2008) to identify JSON structure, and human-in-the-loop (HIL) correction (Zeng et al., 2024a).

4.3 Paradoxical robustness

The erroneous behaviour of the LLM agents could affect the robustness of the model and the approach used. However, the results implied a paradoxical relationship between the LLM-based institutional network model's error-proneness and error-tolerance, which could enhance the understanding of the robustness of multi-agent systems. For instance, with multiple institutional actors joining the system, the chances of erroneous behaviour increase since every single decision-maker has some probability of producing errors. These errors could also be transmitted within the network and affect other agents' decision-making, which, to some extent, corresponds to real-world policy-making. However, with the interaction of multiple agents, no single agent nor their erroneous behaviour had sufficient influence to determine the behaviour of the whole system. The missing output from the research supplier did not lead the system to generate a cascade of unusual behaviour nor did it crash the simulation. The high-level institution's tendency to make balanced decisions also contributed to the error-tolerance of the institutional network. The high-level institution's path-dependent decision-making ensures that the whole system is unlikely to adjust policies drastically. Hence, the incrementalism that derives from the polycentric institutional network structure may help to avoid critical policy failures, which is particularly important in the land system. This could also help to simulate the consequences of widespread distrust between policy actors in large networks (Fischer et al., 2016).

4.4 Contextual coherence does not equal logical consistency

While the agents' performance may reflect certain real-world phenomena within institutional networks, it is essential to address a deeper reflection on the current working mechanisms of LLMs. LLMs are designed to optimize literal contextual coherence, meaning that a vast amount of high-quality, textual data enables the machine to effectively mimic human language by approximating patterns of word (or token) changes (Radford et al., 2018). This creates the illusion that LLMs can think. However, it is crucial to recognize that although logical reasoning when expressed in a language can lead to contextual coherence, the reverse is not necessarily true. In other words, contextual coherence in text might be a necessary condition, but it is not sufficient to produce logical consistency. This raises a caveat: over-anthropomorphizing LLM agents can complicate the evaluation of their outputs. This difficulty arises from both the laborious manual work required to assess the agents' logical consistencies and the logical inconsistencies masked by contextual coherence. In future research, LLMs could be trained using "very strict" language that satisfies the condition that contextual coherence $\implies$ logical consistency, which could ensure that the LLM output is correct. This would be a significant development for LLM-based simulations.

5. Conclusion

We explored the development and application of InsNet-CRAFTY v1.0, a multi-LLM-agent institutional network model with a polycentric structure that is coupled with an agent-based, land system model. By leveraging LLMs to facilitate interactions through textual data, the model enables each modelled entity to pursue unique goals and values that collectively impact the

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

modelled land use system. The results demonstrate that this LLM-enhanced approach is powerful and flexible in modelling institutional actors' behaviours within the land system. However, this novel approach also brings new challenges arising from the limitations of current LLM technology, signifying the need for future research.

Code and data availability. All data and code to run InsNet-CRAFTY version 1.0 are made freely available online via Zenodo (https://doi.org/10.5281/zenodo.13944650, Zeng et al., 2024c; https://doi.org/10.5281/zenodo.13356487, Zeng et al., 2024d)

Author contributions. YZ, CB, and MR contributed to developing the model concept. YZ designed the work, developed the new model and code, and conducted the formal analysis. YZ, CB, JR, and TS wrote the original draft. YZ and MB performed the visualization. CB and MR managed the project administration, while MR acquired funding and supervised the research. All authors reviewed, edited, and validated the final work.

Competing Interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work was supported by the Helmholtz Excellence Recruiting Initiative, Climate Mitigation and Bioeconomy Pathways for Sustainable Forestry Climb-forest (grant no. 101059888) and Co-designing Holistic Forest-based Policy Pathways for Climate Change Mitigation (grant no. 101056755) projects.

Appendix A

Figure A1 illustrates the model processes, segmented into three distinct sections. The land use section (blue) encompasses all processes directly related to changes in land use. The LLM agent section (green) consists of the activities performed by LLM agents. The operational institution is a hybrid agent, integrating rule-based decision-making processes (yellow) and LLM-driven procedures (procedures 4 and 21).

This hybrid approach aligns with the dual nature of organizational routines and non-routine actions, as extensively analysed by Simon in his seminal work, *Administrative Behavior* (Simon, 2013). Organizational routines are recurring actions embedded in an organization's culture, ensuring consistency and efficiency. In contrast, non-routine actions are spontaneous and designed to address unique, unpredictable situations. Both are crucial for effective organizational functioning. The rule-based components correspond to organizational routines, ensuring strict adherence to operational protocols, while the LLM component allows for creative, sometimes imperfect, responses.

In InsNet-CRAFTY, the LLM-related functionalities of the agents are written in Python, while the rule-based processes and CRAFTY are coded in Java. The sub-models written in the two

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

programming languages are connected through a client-server architecture. For a comprehensive description of the rule-based processes within the operational institution (steps 6-14), refer to Zeng et al., (2024b). The explanations of the processes within each section are provided below.

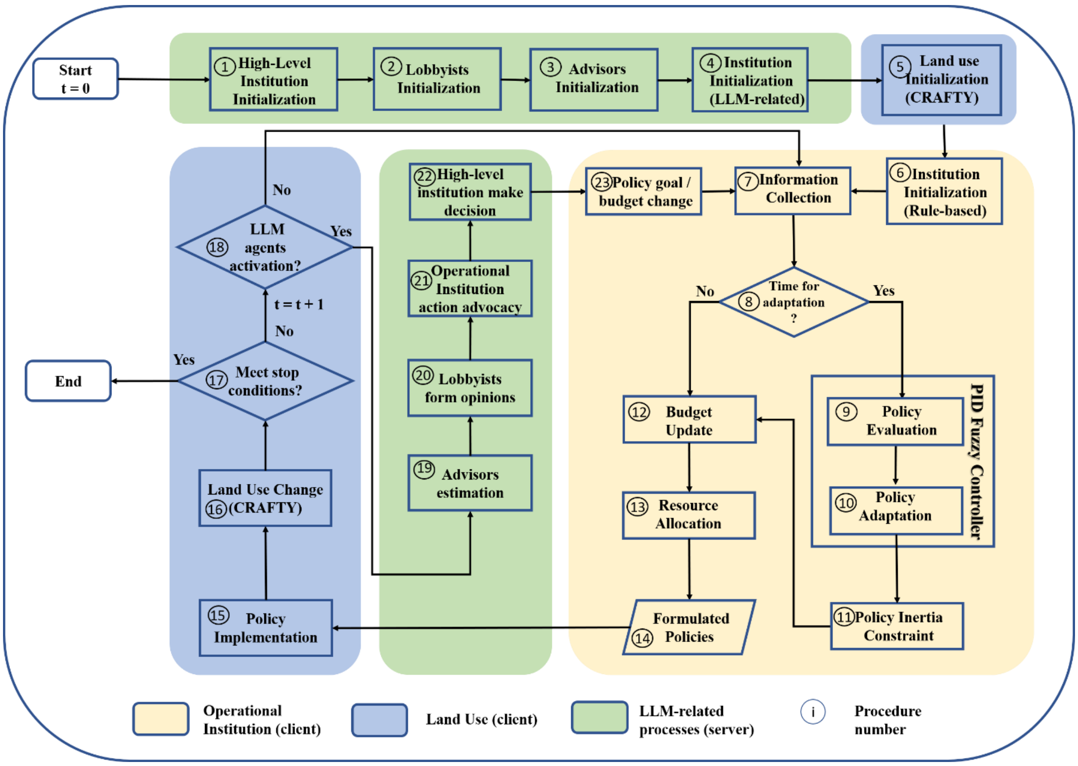

Figure A1: Model processes of InsNet-CRAFTY v1.0

Step 1 – 4: Launching the server end and initialising the LLM agents. This includes creating a server object that listens to requests from the client end and instantiating the agent class by initializing the model names of the large language models, API keys, prompt templates of the LLM agents, and agent-specific work-flows. The optional narrative injector is not displayed here.

Step 5: Launching the client end and initialising the CRAFTY land use model. Key procedures include initializing the distribution of capitals and agent functional types, i.e., AFTs (Brown et al., 2019; Murray-Rust et al., 2014).

Steps 6 – 14: Rule-based policy adaptation of the operational institutions. Step 6 includes the initialization of the operational institutions' rule-based components, initial policy goals and accessible policy instruments. At step 7, each operational institution collects information from the land use model. Step 8 determines if it is time to trigger the policy adaptation processes. If Step 8 outputs true, the operational institution starts to evaluate the current policy's performance using a

PID (Proportional-Integral-Derivative) mechanism and calculate the needed policy intervention intensity using a fuzzy logic controller (Zeng et al., 2024b), which can convert experts' knowledge into computer-comprehensible rules to automate the decision-making (step 9 and 10). Step 11 is a normalised non-monetary constraint restricting the policy change. Steps 12 and 13 further tailor the policy change to satisfy the budgetary constraints. Step 14 is the resultant policy adaptation.

Step 15: Policy implementation. Implement the policy by changing corresponding variables in the land use model.

Step 16: Land use change updating. Run the land use model under the influence of policy interventions. This produces responses of the land use model in terms of land use type distributions and ecosystem services' demand and supply.

Step 17: Termination check. Check if it is time to terminate the whole simulation.

Step 18: LLM interaction check. Check if it is time to trigger the LLM agents. If false, go back to step 7; otherwise, go to step 19 by sending a request and the updated land use data to the server for the policy goals as well as budget allocation formulated through the LLM agents.

Step 19 – 22: LLM agent activation. Activate the LLM agents on the server end to obtain the output of each of them. The narrative injector outputs the updated narratives (optional); the research supplier provides the textual interpretation of the numerical results collected from the land use model; The lobbyists construct their arguments for their benefits; the operational institution's LLM module can also generate arguments to propose financial support and proper policy goal adjustments; the high-level institution receives all the information to form the final decision in terms of budget allocation and policy goal adjustments, which in turn influence the behaviour of the operational institutions.

Step 23: Sending back the updated policy goals and budget allocation to the operational institutions, based on which the operational institutions adjusted their policy-making.

Appendix B

Table B1: Prompt template of the research supplier

You are an AI assistant in policy-making that can write Python code to analyze data.
You are responsible for debugging your own code using the provided tools.
Now, analyse the data in the CSV file in the way you think appropriate.
You can reference the following instructions to conduct your analysis step by step.
Step-by-step data analysis instructions:
1. **Load the Data**
  - Load the CSV file into a DataFrame.
  - Display the first few rows of the DataFrame to understand the structure and the types of data included.

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

- Check for missing values or inconsistencies in the data.
2. **Initial Data Inspection**
- Use descriptive statistics (like `data.describe()`) to get an overview of the numerical features.
- Plot histograms or box plots for each numerical feature to understand the distribution and spot any outliers.
3. **Detailed Analysis of Specific Features**
- **Meat Production Analysis**:
  - Plot time-series graphs for meat demand, meat supply, and policy goals for meat production.
  - Analyze the trends and gaps between demand, supply, and policy goals.
- **Protected Area Analysis**:
  - Plot the protected area ratio over time alongside the policy goals for the protected area ratio.
  - Identify any discrepancies between policy goals and actual outcomes.
4. **Budget Allocation Analysis**
- Create line plots to visualize the budget allocations for meat production and protected areas over time.
- Compare these allocations to see if the budget is aligned with the goals and outcomes.
5. **Evaluate Correlations and Causations**
- Investigate correlations between different variables using scatter plots or correlation matrices.
- Consider potential causative factors that could explain trends observed in the data.
6. **Summarize Findings**
- Summarize key insights into specific bullet points from the data analysis.

Table B2: The prompt template of the environmental NGO

You are a representative of an environmental NGO that is concerned with environmental protection and climate change.
Based on the information given below and the identity of your role, generate some bullet points to lobby the high-level public policy institution to prioritise nature conservation.
The given information: {given_information}

Table B3: Prompt template of the land user association. A specific role – a representative of the meat production industry – is given to the agent to enable it to focus on a concrete topic.

You are a representative of the meat production industry, who cares about the benefits of the industry.

Based on the information given below and the identity of your role, generate some bullet points to lobby the high-level public policy institution to prioritise meat industry development.

The given information: {given_information}

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

Table B4: The prompt template of the agricultural institution

As a policy-maker specializing in agriculture, you oversee initiatives critical to your region's food security, farmer livelihood, and financial well-being. Currently, you're focusing on meat production, a sector facing significant challenges due to changing market demands. Your role is to propose a set of compelling and concise bullet points to the high-level institution, seeking increased priorities and financial support for meat production. Consider the economic impact and social implications. Specifically, you should prompt the high-level institution to make reasonable policy goals that align with budget allocation. Use the data in the CSV file provided to argue your case effectively.

Table B5: The prompt template of the environmental institution

As a policy-maker specializing in environmental protection, you oversee initiatives critical to nature conservation, biodiversity, and pollution reduction including the Net-zero targets in your region. Currently, you're focusing on the expansion of protected areas, a sector facing significant challenges due to biodiversity loss. Your role is to propose a set of compelling and concise bullet points to the high-level institution, seeking increased priorities and financial support for protected area establishment. Specifically, you should prompt the high-level institution to make reasonable policy goals and budget allocation. Use the data in the CSV file provided to argue your case effectively.

Table B6: The prompt template of the law consultant

You are a law consultant giving advice to a high-level public policy institution that is responsible for making public policies regarding agricultural production and environmental protection. Use the provided context about relevant policies, laws, regulations, etc., only to form your advice to ensure the high-level institution makes policies legally. (if you don't know the answer in the given context, just say you don't know): <context> {context} </context> Question: {input}

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

Table B7: Prompt template of the high-level institution

Simulation Role: You are a high-ranking policymaker in charge of overseeing operational institutions within the land system.

Key Actions:

Budget Allocation: Allocate the financial resources between the Agricultural and Environmental Institutions. This directly affects their operational capabilities and initiatives.

Policy Goal Adjustment: Adjusting policy goals appropriately for each institution.

Objective:

Strategically guiding operational institutions, including Agricultural and Environmental Institutions; harmoniously balancing the interests of diverse stakeholders.

Input information:

1) Input from Agricultural institution: {AgriInstInput}
2) Input from Environmental institution: {EnviInstInput}
3) Input from Environmental NGO: {NGOInput}
4) Input from Land user association: {landUserInput}
5) Input from the environment: {narrIput}
6) Input from research suppliers: {researchInput}
7) Historical information: {history}
8) Law consultant: {lawInfo}

Decision-Making Guidance:

Be explicit about your role as a policymaker and your impact on operational institutions.

Make informed decisions by thoroughly analyzing inputs from all stakeholders.

Reflect on historical information to inform decisions.

Ensure that your actions and decisions are logical, well-reasoned, and transparent.

Before giving your final decision, provide a step-by-step rationale for each decision, showing how it aligns to balance stakeholder interests and ensure the feasibility of policy adjustments.

The rationale should support you in quantifying the planned changes in each operational institution's budget and policy goal using percentages.

Note:

The long-term goals have already been specified, your tasks are dynamically conducting reasonable modifications to the goals and providing feasible budget allocation to support the achievement of the goals.

Output requirements:

1. Output your step-by-step reasoning here: including stakeholder input analysis, budget allocation analysis, and policy goal adjustment analysis.
2. Format your quantified policy adjustments using JSON. Your output should be a clean JSON without anything beyond. An example is as follows:

```
{
   "Budget Allocation": {
      "Agricultural Institution": using a positive integer to indicate the percentage of budget allocation here,
      "Environmental Institution": using a positive integer to indicate the percentage of budget allocation here
   },
   "Policy Goal": {
```

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

"Agricultural Institution": using an integer to indicate the percentage of policy goal change here; positive integers indicate the percentage of increase in the current policy goal, while negative ones mean decreasing the current policy goal; 0 means remaining the current policy goal unchanged,
"Environmental Institution": using an integer to indicate the percentage of policy goal change here; positive integers indicate the percentage of increase in the current policy goal, while negative ones mean decreasing the current policy goal; 0 means remaining the current policy goal unchanged
}
}

Appendix C

One can consider the high-level institution as a controller over the operational institutions, which in turn impose their control over the land use model. As previously stated, the operational institutional agents are hybrid. They incorporate both a LLM component to interact with other LLM agents and rule-based behaviour to interact with the programmed land use model. We use the endogenous institutional model described in Zeng et al. (2024b) to simulate the rule-based behaviour of the operational institutional agents. We first describe how the operational institutions' non-LLM modules work, and then introduce how the high-level institution's influence comes into play.

1. Operational institution

1.1 Policy goal definition

The first step to model an operational institution's behaviour is to define a policy goal, which can be represented by a three-dimensional vector:

$$\mathbf{G^i} = [T_s^i, T_e^i, Q^i] \tag{C1}$$

meaning operational institution i's policy goal consists of T_s^i the time when the policy starts, T_e^i the time when the policy ends, and Q^i the quantity of an ecosystem outcome a policy is meant to change during the time from T_s^i to T_e^i. For instance, if we only have two operational institutions, e.g., an environmental institution and an agricultural institution, the possible values of i can only be 1 or 2.

1.2 Policy evaluation and adaptation

The operational institutions estimate their policy effectiveness using Eq. (C2):

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

$$\mathrm{E}_{\mathrm{t_n}} = \frac{1}{k} \sum_{m=n-k}^{n} \frac{Q^i - o_{t_m}^i}{|Q^i|} \tag{C2}$$

where $\mathrm{t_n}$ represents the specific time at which the institution evaluates the goal-outcome error $\mathrm{E}_{\mathrm{t_n}}$; $o_{t_m}^i$ is the outcome intended to be adjusted by institution i at the time t_m; k is the time interval of interest.

Let F denote the function of a fuzzy logic controller (FLC) and $F(E)$ indicate policy variation. The constrained policy variation A_{t+1}^i at $t+1$ is calculated as

$$A_{t+1}^i = sign\big(\mathrm{F}(E)\big) \times \min\big(|\mathrm{F(E)}|, N^i\big) \tag{C3}$$

The above equation means that the absolute value of policy variation within one iteration should be no greater than policy inertia constraint N^i. The sign function outputs the sign (+1, 0, or -1) of its input.

A_{t+1}^i is accumulated to form a policy modifier denoted as M_{t+1}^i, as shown in Eq. (C4).

$$M_{t+1}^i = M_t^i + A_{t+1}^i \tag{C4}$$

The policy variation is normalised and used with a fixed step size for iterative policy adaptation. The policy modifier is a coefficient of the step size. As shown in Eq. (C5), η^i is the step size, and V_{t+1}^i is the modified policy intervention for the $(\mathrm{t}+1)$-th iteration.

$$V_{t+1}^i = \eta^i \times M_{t+1}^i \tag{C5}$$

The budget update process monitors the institution's income and expenditure whenever a policy is implemented. This assumes that policy interventions can be quantitatively measured, with their absolute values being positively correlated with the budget required by the institution to implement the policy. In Eq. (C6), f represents a monotonic function that maps the absolute value of a policy intervention V_{t+1}^i to resource R_{t+1}^i needed to carry out this policy. In this model, only subsidization and the establishment of new protected areas require budget allocations; the costs associated with taxation are not included.

$$R_{t+1}^i = f\left(\left|V_{t+1}^i\right|\right) \tag{C6}$$

The actual policy intervention under the budgetary constraint is

$$V^i \leftarrow \mathrm{sign}(V^i) \times f^{-1}\left(min\left(R_{t+1}^i, B^i\right)\right) \tag{C7}$$

The budget of operational institution i should be updated via operation (C8):

$$B^i \leftarrow \max(B^i - R_{t+1}^i, 0) \tag{C8}$$

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

The implemented policies are supposed to influence land users' behaviour. In CRAFTY (Murray-Rust et al., 2014), land users are categorized into an array of AFTs (Agent Functional Types), each of which can provide multiple ecosystem services. AFTs differ in their capabilities of using a diversity of capitals within land. The AFTs compete for land in the pursuit of benefit, which in turn influences the whole system's ecosystem service supply.

1.3 Policy implementation

In a rasterized map, the competitiveness of an AFT under the influence of economic policies (such as subsidies and taxes) can be calculated as follows:

$$c_{xy} = \sum_{S} (p_S (\sum_{i} V_{ECON}^{iS} + m_S)) \quad \text{(C9)}$$

where c_{xy} denotes the competitiveness of a land use agent at the land cell (x, y); S is the ecosystem service the land user produces; p_S is the total production of S within the land cell; V_{ECON}^{iS} is the institution i's economic policy that targets ecosystem service S; m_S is marginal utility brought by ecosystem service S.

The environmental institution identifies the top N unprotected land cells within the model based on the richness of a chosen set of capitals requiring conservation. Here, two natural capitals defined in the CRAFTY-EU (Brown et al., 2019), i.e., forest and grassland productivity, are used to determine if a land cell needs protection. The value of N at each stage is determined using the previously mentioned fuzzy controller method. Typically, if there is a significant gap between the PA target and the current PA coverage, the value of N would be increased. Certain products cannot be produced on the protected land cells. Therefore, the competitiveness of an AFT on protected land cells can be calculated as:

$$c_{xy} = \sum_{S} (w_S p_S (\sum_{i} V_{t+1}^{iS} + m_S)) \quad \text{(C10)}$$

where w_S represents an element of a vector w whose elements equal either one or zero, which defines if a type of ecosystem service is allowed to be produced in PAs. The CRAFTY-EU model considers seven types of ecosystem service (including meat, crops, habitat diversity, timber, carbon, urban, and recreation). In the current model setting, it is assumed that only habitat diversity is allowed to be improved by the AFTs PAs, reflecting a strict restriction on ecosystem service production.

2. High-level institution

To let the model form a self-sustained system, it is assumed that the total budget obtained by the high-level institution is related to the total production of the ecosystem services, corresponding to the fact that governmental incomes are mainly from the gross domestic product (GDP).

$$B_t^{total} = \alpha \sum_{S} P_{S,t} \quad \text{(C11)}$$

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

where B_t^{total} means the total budget the high-level institution can allocate between the operational institutions at t; $P_{S,t}$ represent the total production of ecosystem service S across all AFTs at time t; α is a coefficient that indicates the proportion of the total budget to total ecosystem service production.

The budget gain $\triangle b_{i,t}$ of operational institution i at time t is calculated as

$$\triangle b_{i,t} = \beta_i B_t^{total} \quad (C12)$$

where β_i is the percentage controlled by the high-level institution. Hence, the budget of operational institution i should be updated:

$$B^i \leftarrow B^i + \triangle b_{i,t} \quad (C13)$$

Whenever the high-level institution adjusts operational institution i's policy goal by a percentage $\triangle q^i$, the policy goal is updated as

$$Q^i \leftarrow Q^i(1 + \triangle q^i) \quad (C14)$$

Operation (C7) indicates that operational institutions cannot consume resources more than their budget. However, the equation does imply that the budget can be insufficient for implementing a policy. We use the difference between operational institution i's budget and the needed resources to calculate the budget surplus at time t using Eq. (C15). Therefore, the budget surplus can be either positive or negative.

$$SUR_{i,t} = B^i - R_t^i \quad (C15)$$

3. Numerical settings

Table C1 – C3 show the numerical settings related to the policy-making processes of the operational institutions.

Table C1: The settings of the operational institutions and the high-level institution

Institution attributes	**Settings**
Unique ID	"Agricultural_Institution"
Policies	"Meat_economic"
Information	Annual meat supply and demand, budget surplus.
Budget	Allocated by the high-level institution based on total ecosystem service production annually.
Decision rules	"Economic"
Policy attributes	
Unique ID	"Meat_economic"
Target service	"Meat"

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

Policy Type	"Economic" (see Table S2)
Step size η[1]	1000000
Inertia constraint	1.0
Initial policy goal	120% initial meat production
Time lag	2
Policy-resource function	$R = f(\|V\|) = max(V, 0)$ (Note: Only if $V > 0$, does the institution consume budget, and the budget use equals the subsidy.)
Institution attributes	**Settings**
Unique ID	"Environmental_Institution"
Policies	"Protected_areas"
Information	Protected area ratio
Budget	Allocated by the high-level institution based on total ecosystem service production annually.
Decision rules	"Protection" (see Table S3)
Policy attributes	
Unique ID	"Protected_areas"
Target service	"Protected areas"
Policy Type	"Protection"
Step size η[2]	1.0
Initial policy goal	10% of total land
Initial guess	10000
Time lag	2
Timer Adapting	Equal to the time lag False
Policy-resource function	$R = f(\|V\|) = 1000V$ (Note: *V* indicates the number of land cells that need to be protected, and it is assumed that each new protected cell consumes 1000 units of budget. The value is set for making the budget consumptions of the two operational institutions comparable.)
Institution attributes	**Settings**
α	0.01 (Note: The high-level institution uses 0.01 times the total ecosystem production of the modelled system as the total budget that can be allocated between the two operational institutions)

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

Table C2: Parameterisation of the fuzzy decision rules labelled as “Economic”, using FLC language defined in the IEC 61131-7 (IEC 61131-7, 2024)

```
FUNCTION_BLOCK Economic
VAR_INPUT
        gap: REAL;
END_VAR
VAR_OUTPUT
        Intervention : REAL;
END_VAR
FUZZIFY gap
        TERM nhigh := (-0.5,1) (-0.3,0);
        TERM nmild := (-0.5,0) (-0.3,1) (-0.1,0);
        TERM nlight := (-0.3,0) (-0.1,1) (0,0);
        TERM neutral := (-0.05,0) (0,1) (0.05,0);
        TERM plight := (0, 0) (0.1, 1) (0.3,0);
        TERM pmild := (0.1,0) (0.3,1) (0.5,0);
        TERM phigh := (0.3, 0) (0.5, 1);
END_FUZZIFY
DEFUZZIFY intervention
        TERM nhigh := (-0.2,1) (-0.1,0);
        TERM nmild := (-0.15,0) (-0.05,1) (0,0);
        TERM neutral := (-0.02,0) (0,1) (0.02,0);
        TERM pmild := (0,0) (0.05,1) (0.15,0);
        TERM phigh := (0.1,0) (0.2,1);

        METHOD : COG;
        DEFAULT := 0;
END_DEFUZZIFY
RULEBLOCK No1
        AND : MIN;
        ACT : MIN;
        ACCU : MAX;

        RULE 1 : IF gap IS nhigh  THEN intervention IS nhigh;
        RULE 2 : IF gap IS nmild  THEN intervention IS nmild;
        RULE 3 : IF gap IS nlight THEN intervention IS neutral;
        RULE 4 : IF gap IS neutral THEN intervention IS neutral;
        RULE 5 : IF gap IS plight THEN intervention IS neutral;
        RULE 6 : IF gap IS pmild THEN intervention IS pmild;
        RULE 7 : IF gap IS phigh THEN intervention IS phigh;
END_RULEBLOCK
END_FUNCTION_BLOCK
```

Table C3: Parameterisation of fuzzy decision rules labelled as “Protection”, FLC language defined in the IEC 61131-7 (IEC 61131-7, 2024)

```
FUNCTION_BLOCK Protection
VAR_INPUT
        gap: REAL;
END_VAR
```

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024


```
VAR_OUTPUT
        intervention : REAL;
END_VAR
FUZZIFY gap
        TERM plow := (0,1) (0.15,0);
        TERM plight := (0.025, 0) (0.175, 1) (0.325,0);
        TERM pmild := (0.175,0) (0.325,1) (0.45,0);
        TERM phigh := (0.325, 0) (0.45, 1);
END_FUZZIFY
DEFUZZIFY intervention
        TERM neutral := (0,1) (0.075,0);
        TERM plight := (0.025,0) (0.075,1) (0.125,0);
        TERM pmild := (0.075,0) (0.125,1) (0.175,0);
        TERM phigh := (0.125,0) (0.2,1);

        METHOD : COG;
        DEFAULT := 0;
END_DEFUZZIFY
RULEBLOCK No1
        AND : MIN;
        ACT : MIN;
        ACCU : MAX;

        RULE 0 : IF gap IS plow THEN intervention IS neutral;
        RULE 1 : IF gap IS plight THEN intervention IS plight;
        RULE 2 : IF gap IS pmild THEN intervention IS pmild;
        RULE 3 : IF gap IS phigh THEN intervention IS phigh;
END_RULEBLOCK
END_FUNCTION_BLOCK
```

Reference

Ariti, A. T., van Vliet, J., and Verburg, P. H.: The role of institutional actors and their interactions in the land use policy making process in Ethiopia, J. Environ. Manage., 237, 235–246, https://doi.org/10.1016/j.jenvman.2019.02.059, 2019.

Bainbridge, A., Troppe, T., and Bartley, J.: Responding to research evidence in Parliament: A case study on selective education policy, Rev. Educ., 10, e3335, https://doi.org/10.1002/rev3.3335, 2022.

Barnett, M. N. and Finnemore, M.: The politics, power, and pathologies of international organizations, Int. Organ., 53, 699–732, https://doi.org/10.1162/002081899551048,1999.

Broussard, A., Dahdouh-Guebas, F., and Hugé, J.: Diversity of perspectives in biodiversity conservation: A case study of port land use in Antwerp and Rotterdam, J. Environ. Manage., 341, 117937, https://doi.org/10.1016/j.jenvman.2023.117937, 2023.

Brown, C., Seo, B., and Rounsevell, M.: Societal breakdown as an emergent property of large-scale behavioural models of land use change, Earth Syst. Dyn., 10, 809–845, https://doi.org/10.5194/esd-10-809-2019 , 2019.

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

Cairney, P., Heikkila, T., and Wood, M.: Making policy in a complex world, Cambridge University Press, 2019.

Calvin, K. V., Snyder, A., Zhao, X., and Wise, M.: Modeling land use and land cover change: using a hindcast to estimate economic parameters in gcamland v2.0, Geosci. Model Dev., 15, 429–447, https://doi.org/10.5194/gmd-15-429-2022, 2022.

Chen, M., Vernon, C. R., Huang, M., Calvin, K. V., and Kraucunas, I. P.: Calibration and analysis of the uncertainty in downscaling global land use and land cover projections from GCAM using Demeter (v1.0.0), Geosci. Model Dev., 12, 1753–1764, https://doi.org/10.5194/gmd-12-1753-2019, 2019.

Dakin, R. and Ryder, T. B.: Reciprocity and behavioral heterogeneity govern the stability of social networks, Proc. Natl. Acad. Sci., 117, 2993–2999, https://doi.org/10.1073/pnas.1913284117, 2020.

Davidson, M. R., Filatova, T., Peng, W., Verbeek, L., and Kucuksayacigil, F.: Simulating institutional heterogeneity in sustainability science, Proc. Natl. Acad. Sci., 121, e2215674121, https://doi.org/10.1073/pnas.2215674121,2024.

Díez-Echavarría, L., Villegas-Palacio, C., Arango-Aramburo, S., and Ezzine-de-Blas, D.: Decoupling in governance: the land governance network in a region of the Colombian Andes, Land Use Policy, 133, 106880, https://doi.org/10.1016/j.landusepol.2023.106880, 2023.

Djekic, I.: Environmental Impact of Meat Industry – Current Status and Future Perspectives, Procedia Food Sci., 5, 61–64, https://doi.org/10.1016/j.profoo.2015.09.025, 2015.

Duffy, C., Prudhomme, R., Duffy, B., Gibbons, J., O'Donoghue, C., Ryan, M., and Styles, D.: GOBLIN version 1.0: a land balance model to identify national agriculture and land use pathways to climate neutrality via backcasting, Geosci. Model Dev., 15, 2239–2264, https://doi.org/10.5194/gmd-15-2239-2022, 2022.

Ekholm, T., Freistetter, N.-C., Rautiainen, A., and Thölix, L.: CLASH – Climate-responsive Land Allocation model with carbon Storage and Harvests, Geosci. Model Dev., 17, 3041–3062, https://doi.org/10.5194/gmd-17-3041-2024, 2024.

Fan, W., Ding, Y., Ning, L., Wang, S., Li, H., Yin, D., Chua, T.-S., and Li, Q.: A Survey on RAG Meeting LLMs: Towards Retrieval-Augmented Large Language Models, arXiv preprint, http://arxiv.org/abs/2405.06211, 2024.

Feindt, P. H.: Policy-Learning and Environmental Policy Integration in the Common Agricultural Policy, 1973–2003, Public Adm., 88, 296–314, https://doi.org/10.1111/j.1467-9299.2010.01833.x, 2010.

Fischer, M., Ingold, K., Sciarini, P., and Varone, F.: Dealing with bad guys: Actor-and process-level determinants of the "devil shift" in policy making, J. Public Policy, 36, 309–334, https://doi.org/10.1017/S0143814X15000021, 2016.

Fischer, P., Schulz-Hardt, S., and Frey, D.: Selective exposure and information quantity: how different information quantities moderate decision makers' preference for consistent and inconsistent information., J. Pers. Soc. Psychol., 94, 231, https://doi.org/10.1037/0022-3514.94.2.94.2.231, 2008.

Gigerenzer, G. and Goldstein, D. G.: Reasoning the fast and frugal way: models of bounded rationality., Psychol. Rev., 103, 650, https://doi.org/10.1037/0033-295x.103.4.650, 1996.

González, V. B.: Modelling adaptation strategies for Swedish forestry under climate and global change, University of Edinburgh, 2016.

Guo, J., Li, F. Y., Tuvshintogtokh, I., Niu, J., Li, H., Shen, B., and Wang, Y.: Past dynamics and future prediction of the impacts of land use cover change and climate change on landscape ecological risk across the Mongolian plateau, J. Environ. Manage., 355, 120365, https://doi.org/10.1016/j.jenvman.2024.120365, 2024.

Hallerberg, M. and Von Hagen, J.: Electoral institutions, cabinet negotiations, and budget deficits in the European Union, National Bureau of Economic Research Cambridge, Mass., USA, 1997.

Hong, S., Zhuge, M., Chen, J., Zheng, X., Cheng, Y., Zhang, C., Wang, J., Wang, Z., Yau, S. K. S., Lin, Z., Zhou, L., Ran, C., Xiao, L., Wu, C., and Schmidhuber, J.: MetaGPT: Meta Programming for A Multi-Agent Collaborative Framework, arXiv preprint, http://arxiv.org/abs/2308.00352, 2023.

IEC 61131-7: https://plcopen.org/iec-61131-7, last access: 22 August 2024.

Ji, Z., Yu, T., Xu, Y., Lee, N., Ishii, E., and Fung, P.: Towards mitigating LLM hallucination via self reflection, in: Findings of the Association for Computational Linguistics: EMNLP 2023, 1827–1843, https://doi.org/10.18653/v1/2023.findings-emnlp.123, 2023.

Jones, B. D.: Bounded Rationality and Political Science: Lessons from Public Administration and Public Policy, J. Public Adm. Res. Theory, 13, 395–412, https://doi.org/10.1093/jpart/mug028, 2003.

Kimmich, C., Ehlers, M.-H., Kellner, E., Oberlack, C., Thiel, A., and Villamayor-Tomas, S.: Networks of action situations in social–ecological systems: current approaches and potential futures, Sustain. Sci., 18, 1–10, https://doi.org/10.1007/s11625-022-01278-w, 2023.

LangChain: https://python.langchain.com/v0.1/docs/get_started/introduction/, last access: 9 May 2024.

Lawrence, A., Houghton, J., Thomas, J., and Weldon, P.: Where Is the Evidence? Realising the Value of Grey Literature for Public Policy & Practice, A Discussion Paper, Discuss. Pap., 2014.

Lee, J., Kim, S., and Lee, J.: Public vs. Public: Balancing the Competing Public Values of Participatory Budgeting, Public Adm. Q., 46, 39–66, https://doi.org/10.37808/paq.46.1.3, 2022.

Li, Y., Krishnamurthy, R., Raghavan, S., Vaithyanathan, S., and Jagadish, H.: Regular expression learning for information extraction, in: Proceedings of the 2008 conference on empirical methods in natural language processing, 21–30, 2008.

Lindblom, C. E.: The science of muddling through, Read. Manag. Psychol., 4, 117–131, https://doi.org/10.2307/973677, 1989.

Liu, Z., Yao, W., Zhang, J., Yang, L., Liu, Z., Tan, J., Choubey, P. K., Lan, T., Wu, J., Wang, H., Heinecke, S., Xiong, C., and Savarese, S.: AgentLite: A Lightweight Library for Building and Advancing Task-Oriented LLM Agent System, http://arxiv.org/abs/2402.15538, 2024.

Machovina, B., Feeley, K. J., and Ripple, W. J.: Biodiversity conservation: The key is reducing meat consumption, Sci. Total Environ., 536, 419–431, https://doi.org/10.1016/j.scitotenv.2015.07.022, 2015.

McGinnis, M. D.: Networks of Adjacent Action Situations in Polycentric Governance, Policy Stud. J., 39, 51–78, https://doi.org/10.1111/j.1541-0072.2010.00396.x, 2011.

Minaee, S., Mikolov, T., Nikzad, N., Chenaghlu, M., Socher, R., Amatriain, X., and Gao, J.: Large Language Models: A Survey, arXiv preprint, http://arxiv.org/abs/2402.06196, 2024.

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

Modarressi, A., Imani, A., Fayyaz, M., and Schütze, H.: Ret-llm: Towards a general read-write memory for large language models, arXiv preprint, https://doi.org/10.48550/arXiv.2305.14322, 2023.

Murray-Rust, D., Brown, C., van Vliet, J., Alam, S. J., Robinson, D. T., Verburg, P. H., and Rounsevell, M.: Combining agent functional types, capitals and services to model land use dynamics, Environ. Model. Softw., 59, 187–201, https://doi.org/10.1016/j.envsoft.2014.05.019, 2014.

Nair, S. and Howlett, M.: Policy myopia as a source of policy failure: adaptation and policy learning under deep uncertainty, Policy Polit., 45, 103–118, https://doi.org/10.1332/030557316X14788776017743, 2017.

Neri, S. and Ropele, T.: Imperfect information, real-time data and monetary policy in the euro area, Econ. J., 122, 651–674, https://doi.org/10.1111/j.1468-0297.2011.02488.x, 2012.

Park, J. S., O'Brien, J. C., Cai, C. J., Morris, M. R., Liang, P., and Bernstein, M. S.: Generative Agents: Interactive Simulacra of Human Behavior, arXiv preprint, http://arxiv.org/abs/2304.03442, 2023.

Perkins, O., Kasoar, M., Voulgarakis, A., Smith, C., Mistry, J., and Millington, J. D. A.: A global behavioural model of human fire use and management: WHAM! v1.0, Geosci. Model Dev., 17, 3993–4016, https://doi.org/10.5194/gmd-17-3993-2024, 2024.

Perković, G., Drobnjak, A., and Botički, I.: Hallucinations in LLMs: Understanding and Addressing Challenges, in: 2024 47th MIPRO ICT and Electronics Convention (MIPRO), 2084–2088, https://doi.org/10.1109/MIPRO60963.2024.10569238, 2024.

Petrovic, Z., Djordjevic, V., Milicevic, D., Nastasijevic, I., and Parunovic, N.: Meat Production and Consumption: Environmental Consequences, Procedia Food Sci., 5, 235–238, https://doi.org/10.1016/j.profoo.2015.09.041, 2015.

Qi, X., Wang, R. Y., Li, J., Zhang, T., Liu, L., and He, Y.: Ensuring food security with lower environmental costs under intensive agricultural land use patterns: A case study from China, J. Environ. Manage., 213, 329–340, https://doi.org/10.1016/j.jenvman.2018.02.048, 2018.

Qian, C., Liu, W., Liu, H., Chen, N., Dang, Y., Li, J., Yang, C., Chen, W., Su, Y., Cong, X., Xu, J., Li, D., Liu, Z., and Sun, M.: ChatDev: Communicative Agents for Software Development, arXiv preprint, http://arxiv.org/abs/2307.07924, 2024.

Radford, A., Narasimhan, K., Salimans, T., Sutskever, I., and others: Improving language understanding by generative pre-training, 2018.

Rich, R. F.: Selective utilization of social science related information by federal policy-makers, Inquiry, 12, 239–245, https://www.jstor.org/stable/29770949, 1975.

Simon, H. A.: Administrative behavior, Simon and Schuster, 2013.

Simon, H. A. : Theories of bounded rationality, Decis. Organ., 1, 161–176, 1972.

Sinden, A.: In Defense of Absolutes: Combating the Politics of Power in Enviornmental Law, Iowa Law Rev., 90, 1405, http://dx.doi.org/10.34944/dspace/6615, 2004.

Sumers, T. R., Yao, S., Narasimhan, K., and Griffiths, T. L.: Cognitive Architectures for Language Agents, arXiv preprint, http://arxiv.org/abs/2309.02427, 2024.

Tan, R., Xiong, C., and Kimmich, C.: An agent-situation-based model for networked action situations: Cap-and-trade land policies in China, Land Use Policy, 131, 106743, https://doi.org/10.1016/j.landusepol.2023.106743, 2023.

https://doi.org/10.5194/egusphere-2024-2661
Preprint. Discussion started: 21 October 2024

Tesfaye, M., Kimengsi, J. N., and Giessen, L.: A policy mix for achieving ambitious goals on forest landscape restoration: Analyzing coherence and consistency in Ethiopia forest-related policy, Land Use Policy, 144, 107214, https://doi.org/10.1016/j.landusepol.2024.107214, 2024.

Tonmoy, S. M. T. I., Zaman, S. M. M., Jain, V., Rani, A., Rawte, V., Chadha, A., and Das, A.: A Comprehensive Survey of Hallucination Mitigation Techniques in Large Language Models, arXiv preprint, http://arxiv.org/abs/2401.01313, 2024.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I.: Attention is all you need, Adv. Neural Inf. Process. Syst., 30, 2017.

Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., Chen, Z., Tang, J., Chen, X., Lin, Y., Zhao, W. X., Wei, Z., and Wen, J.: A survey on large language model based autonomous agents, Front. Comput. Sci., 18, 186345, https://doi.org/10.1007/s11704-024-40231-1, 2024.

Wu, Q., Bansal, G., Zhang, J., Wu, Y., Li, B., Zhu, E., Jiang, L., Zhang, X., Zhang, S., Liu, J., Awadallah, A. H., White, R. W., Burger, D., and Wang, C.: AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation, arXiv preprint, http://arxiv.org/abs/2308.08155, 2023.

Yami, M., van Asten, P., Hauser, M., Schut, M., and Pali, P.: Participation without negotiating: Influence of stakeholder power imbalances and engagement models on agricultural policy development in Uganda, Rural Sociol., 84, 390–415, https://doi.org/10.1111/ruso.12229, 2019.

Yang, H., Yue, S., and He, Y.: Auto-GPT for Online Decision Making: Benchmarks and Additional Opinions, arXiv preprint, http://arxiv.org/abs/2306.02224, 2023.

Yao, J.-Y., Ning, K.-P., Liu, Z.-H., Ning, M.-N., and Yuan, L.: LLM Lies: Hallucinations are not Bugs, but Features as Adversarial Examples, arXiv preprint, http://arxiv.org/abs/2310.01469, 2023.

Zeng, Y., Brown, C., Raymond, J., Byari, M., Hotz, R., and Rounsevell, M.: Exploring the opportunities and challenges of using large language models to represent institutional agency in land system modelling, EGUsphere preprint, https://doi.org/10.5194/egusphere-2024-449, 2024a.

Zeng, Y., Raymond, J., Brown, C., Byari, M., and Rounsevell, M. D. A.: Simulating Endogenous Institutional Behaviour and Policy Pathways within the Land System, SSRN preprint, https://doi.org/10.2139/ssrn.4814296, 2024b.

Zeng, Y., Brown, C., Byari, M.,Raymond, J., Schmitt, T., and Rounsevell, M. D. A.: Data for running InsNet-CRAFTY v1.0 (1.0), Zenodo, https://doi.org/10.5281/zenodo.13944650 [data set], 2024c

Zeng, Y., Brown, C., Byari, M.,Raymond, J., Schmitt, T., and Rounsevell, M. D. A.: Code for running InsNet-CRAFTY v1.0 (1.0), Zenodo, https://doi.org/10.5281/zenodo.13356487 [code], 2024d

Zhong, W., Guo, L., Gao, Q., Ye, H., and Wang, Y.: Memorybank: Enhancing large language models with long-term memory, in: Proceedings of the AAAI Conference on Artificial Intelligence, 19724–19731, https://doi.org/10.1609/aaai.v38i17.29946, 2024.

Zhou, X., Li, G., and Liu, Z.: LLM As DBA, arXiv preprint, https://doi.org/10.48550/arXiv.2308.05481, 2023.