# Iseec

## Modelling human–natural systems interactions with implications for twenty-first-century Warming

【关于气候缓解的研究涉及社会应该做什么】Most studies on climate mitigation address the question of what society should do. Such normative approaches have led to aspirational solutions such as achieving carbon neutrality by 2050 1 , reducing short-lived climate pollutants (SLCPs) and extracting up to one trillion tons of CO2 from the air, among others. These prescriptive solutions have been enormously helpful in guiding mitigation policies. 【现有的研究与现实情况相悖】However, the reality is that the growth of CO2 emissions and global warming have continued unabated2 . 【目前的减排目标没有考虑到相互作用、阻碍了系统设计】Part of the problem is that the aspirational mitigation goals do not always fully factor in interactions between human and natural systems. Such interactions create barriers against redesigning energy and economic systems3,4 .

【许多人认为应该采取新方法来考虑人与自然作用的方面，并且新方法是考虑相互作用】Many have concluded that it is now time for a new approach5,6 , which explicitly allows for interactions between natural systems and social systems7–13. 【目前采用新方法来考虑人与自然相互作用的研究例子有哪些，他们是怎么考虑的】Examples of human–natural interactions that have been considered include societal response to perceived climate risks such as providing support for mitigation policies and changing personal behaviours that reduce individual greenhouse gas emissions7 ; the role of finance and other macroeconomic agents in sustaining economic development and shaping the dynamics of transition to low-carbon energy sources9 ; the role of positive tipping points in the societal transformation towards stabilizing climate10; negative feedback on the economics of energy production due to climate risks12; and constraints on carbon emission reductions due to legacy technologies10 and climate change11. Another recent development in coupling human and natural systems is the modelling of the feedbacks between warming, land use and land-cover policy14,15, which has paved the way for an integrated Earth system model16.

【问题研究的模型代表IAM目前的重要作用】The increasing use of integrated assessment models (IAMs)9 has substantially influenced climate mitigation actions proposed in the most recent Intergovernmental Panel on Climate Change (IPCC) Working Group III report17. Most IAMs, 【目前IAM模型的不足：考虑之前的方面太少】however, are still in the early stages of development for fully coupling nature and human systems. As concluded by Peng et al.18, “IAMs have to get real about people”. For example, a recent review19on integrated human–Earth system modelling stated that ‘many potential human–Earth system feedbacks were not explored in these studies’ and ‘none of the studies look at the interactions between energy and climate.’ 【这项研究工作的重点，贡献之处】The present study attempts to fill this gap by focusing on energy–climate interactions

# MDP-Multiagent+water

## Assessing Adaptive Irrigation Impacts on Water Scarcity in Nonstationary Environments—A Multi-Agent Reinforcement Learning Approach

【问题的引出，历史背景：由于1955的计划启发，水资源管理倾向于的工作是xx】Since the launch of the Harvard Water Program in 1955, the field of water resources management has been dedicated to providing practical solutions for complex water systems (Reuss, 2003). 【然而，这方面工作存在阻碍】However, these efforts are hindered by the uncertainty and nonstationarity rooted in both the human (demand) and natural (sup ply) systems (Brown et al., 2015; Cosgrove & Loucks, 2015; Herman et al., 2020). 【目前有一些方法来处理，概括总结方法】Several planning approach es have emerged to address this issue, which can be grouped into three categories: **dynamic planning, risk based planning, and robust planning.** The first category, dynamic planning, seeks a set of decision rules that yields the optimal outcomes (Castelletti et al., 2010; Harrison, 2007; Higgins et al., 2008). Risk-based planning aims at generating solutions under different risk levels and explores tradeoffs between risks and management objectives (Borgomeo et al., 2018; Hall et al., 2020; Lund, 2002), whereas robust planning looks for strategies that are robust (acceptable) across a variety of future conditions (Lempert & Collins, 2007). 【这些方法的共同之处】These categories are not mutually exclusive as several hybrid methods have been proposed. 【举例说明】For instance, Dynamic Adaptive Policy Pathways incorporates robust decision-making into dynamic planning to determine the timing and actions for adaptation (Haasnoot et al., 2013; Kwakkel et al., 2015, 2016). Many-objective Robust Decision Making explores robust planning alternatives, plans that perform well in a set of future conditions, and the tradeoffs among management objectives (including risks) (Kasprzyk et al., 2013; Watson & Kasprzyk, 2017; Yan et al., 2017). Another example is risk-based stochastic programming which enhances dynamic planning by considering risk-aversion to avoid undesirable outcomes (Piantadosi et al., 2008) and learning by doing to improve decision-making adaptively (Hung & Hobbs, 2019).【然而，这些方法存在没有考虑人与自然的回馈过程不足】 Nevertheless, the meth ods adopted in these studies often do not acknowledge the active interactions among human water users (stakeholders) and natural water supply. 【如果没有考虑不足有哪些不良后果】Water systems are inherently complex human-nature systems as stakeholders can interact with each other and the environment. Neglecting or simplifying the dynamics among stakeholders and stakeholders' adaptive behaviors may lead to biased conclusions.

【举例子说明考虑这部分互馈的重要性，好处】Studies have indicated that incorporation of stakeholders objectives (Gold et al., 2019; Hadjimichael et al., 2020; Quinn et al., 2017) and consideration of stakeholders' cognitive beliefs and values (Glynnet al., 2018; Moallemi et al., 2020), are critical for managing complex human-nature systems. 【ABM的技术的简单介绍，引入领域的时间，后续发展】Originat ing from the Artificial Intelligence community, agent-based modeling (ABM) was introduced to the water resources field in the late 2000s (Berglund, 2015). Since then, ABMs have emerged for assessing human systems dynamics and impacts on water systems (e.g., Al-Amin et al., 2018; Berglund, 2015; Ng et al., 2011; Yang et al., 2009). ABM is a distributed, bottom-up planning approach for understanding human impacts on system performance. An agent in an ABM is an object that interacts with other agents and the system of interest (a virtual environment, e.g., water resources models). et al., 2018; Moallemi et al., 2020), are critical for managing complex human-nature systems. Originat ing from the Artificial Intelligence community, agent-based modeling (ABM) was introduced to the water resources field in the late 2000s (Berglund, 2015). Since then, ABMs have emerged for assessing human systems dynamics and impacts on water systems (e.g., Al-Amin et al., 2018; Berglund, 2015; Ng et al., 2011; Yang et al., 2009).【具体分析方法的详细特点】ABM is a distributed, bottom-up planning approach for understanding human impacts on system performance. An agent in an ABM is an object that interacts with other agents and the system of interest (a virtual environment, e.g., water resources models). the collective behavior of the water users in that group, but it may not represent the individual water users' decision-making. 【方法的具体分类】Berglund (2015) summarized earlier ABM development and developed case studies of two distinct agent types for water resources planning: the reactive agents respond to environmental signals based on behavio ral rules, and the active agents pursue strategies to optimize their objectives. The former is also known as a descriptive model, and the latter a normative model (Smith, 1991). Recently, several ABM frameworks have been proposed to address the natural uncertainties and nonstationarity of complex human-natural systems. Giuliani et al. (2016) proposed a normative ABM framework to investigate the coevolution of agricultural water systems under climate change. Hyun et al. (2019) and Yang et al. (2020) applied psychological the ories to simulate farmers’ irrigation decision-making under uncertainty. Al-Amin et al. (2018) coupled a descriptive ABM with a groundwater model to evaluate water restriction programs in a watershed with mul tiple cities under future climate scenarios. Castilla-Rho et al. (2015) developed a similar ABM-groundwater model framework with a focus on participatory modeling. 【具体分类方法各有哪些不足之处】Generally, normative ABMs assume rationality and disregard human cognitive activities (e.g., learning and risk attitude), while descriptive ABMs only simulate encoded behavior rules that may misrepresent human response when the environment changes. 【我们认为引入RL可以解决传统方法的不足】We believe this research gap can be addressed, at least partially, by reinforcement learning (RL) algorithms.

【引出RL方法的特点，介绍】RL is an area of machine learning where an intelligent agent can learn and improve its decision-making to optimize the long-term reward (Sutton & Barto, 1998). 【RL方法的处理的两种问题区别】Depending on the context of the problems, RL meth ods can be either learning algorithms for searching optimal policy (a set of rules that guides an agent's actions in RL's terms) or decision-making models for simulating human adaptive behaviors (Seo & Lee, 2017; Sutton & Barto, 1998). 【具体介绍RL方法处理的问题种类1、2】The former is termed as learning methods, and the latter termed planning methods in the RL community (Sutton, 1992). RL has two essential characteristics: trial-and-error search and delayed reward, which enable agents to adapt their strategies by interacting with the environment. 【RL引入水资源领域的应用有哪些】Since the intro duction of RL to the water resources field, it has been extensively applied for optimal reservoir operations (Castelletti et al., 2010; Dariane & Moradi, 2016; Lee & Labadie, 2007; Madani & Hooshyar, 2014; Rieker & Labadie, 2012), with a few exceptions in dam sizing (Bertoni et al., 2020) and water and natural resource allocations (Bone & Dragićević, 2009; Ni et al., 2014).【然而，RL引入水资源应用的缺陷】 However, despite its increasing popularity, these appli cations only focus on RL's learning aspect for finding optimal policies in a stationary environment.

【我们聚焦的特点，提出了RL-ABM结合的框架考虑】Focusing on the RL's planning aspect, this paper proposes an RL-ABM framework that equips agents (i.e., the agriculture water users in the case study) with the ability to adapt to a changing water system. 【我们的RL框架中基本考虑了哪些】The agents' decisions are the water quantity requests submitted to the water resources administration, and the water system is assumed under the impacts of climate change and growing water consumption.【RL方法的好处，适宜于陈列问题】 Rooted in psychology and cognitive science, the RL algorithm is advantageous, compared to the aforementioned methods, in simulating the constant changes in human beliefs and strategies through interactions with the environment. This feature is critical for the simulation of human reactions in nonstationary systems. 【补充分析RL具体的细节好处】Moreover, RL's parameters are relevant to human cognitive activities and can be used to characterize agents’ behaviors (i.e., water diversion patterns in response to signals of environmental changes from the water system).

【本文的框架应用的区域是Colorado River】The RL-ABM framework is applied to the Colorado River Basin (CRB), United States (US), as an illustrative case study. 【具体介绍研究区】The CRB is one of the most critical water sources in the Western US and Mexico and is facing increasing water stress due to recent droughts and the warming climate (US Bureau of Reclamation, 2012). 【环境对应的模拟】The Colorado River Simulation System (CRSS), a water resources management model for CRB developed by the US Bureau of Reclamation (USBR) for reservoir operations and policy evaluations (US Bureau of Reclamation, 2007a; Zagona et al., 2001), is adopted in the case study as the virtual environment to be coupled with the RL-ABM for the assessment of water system response to dynamic agriculture water demands. 【agent可以代表的主体】The RL-ABM agents (agriculture water users in the CRSS) can be a single farm, an irrigation ditch, an irrigation district, American tribal water users, or a group of farming entities. Therefore, an agent represents only the collective behavior of the water users in that group, but it may not represent the individual water users' decision-making.

【总结，论文的贡献点】In summary, this paper's contribution includes: (a) a modeling approach for studying human adaptation impacts on nonstationary water resources systems with a focus on the human cognitive aspect, and (b) an RL algorithm for simulating agriculture water users’ adaptive water consumption decisions that incorporate additional water availability information. 【然而，RL-ABM的参数可以塑造意识过程提供主体对环境反馈】Moreover, RL-ABM parameters can better characterize an agent's cognitive processes comparing to Hyun et al. (2019), which can provide information about agents’ reactions to environmental changes.

【文章结构书写顺序的安排】The remainder of this paper is organized as follows. Section 2 describes the modeling framework for agriculture water users’ adaptive policy and the ABM-CRSS coupling. Section 3 introduces the case study. The results are shown in Section 4, followed by the discussion in Section 5. Finally, we present our conclusions and final remarks in Section 6.

# Pnas文章

## Reinforcement learning–based adaptive strategies for climate change adaptation: An application for coastal flood risk management

【气候变化中需要做决策，决策过程很复杂】Effectively adapting to climate change necessitates the development of robust policies and public investments within systems that are complex, interlinked, and characterized by manifold uncertainties. 【做决策过程中具体为什么复杂】Planners therefore need to be able to adapt their plans to the changing reality ( 1 ). Dynamically responsive, flexible adaptation strategies generally outperform rigid, fixed strategies by incorporating evolving knowledge about risks and uncertainties. 【决策投资过程中具体做的步骤】First, phased investments can be planned, facilitating relatively low-cost initial actions. Second, dynamic investments enable adjusting actions or plans in response to unexpected future states, forestalling catastrophic failures. Third, possible future actions can be considered in current decision-making to avoid overestimating lifetime risks

【决策的适应性框架概念】Flexible adaptation frameworks are diversely referred to as “adaptation pathways” ( 2 , 3 ), “dynamic adaptation” ( 4 , 5 ), or “real options analysis” ( 6 , 7 ). 【描述该概念的方法有哪些】Several analytical approaches have been developed for modeling these policy frameworks ( 8 – 10 ) and applied to evaluate the benefits and costs of adaptation measures, but they do not account for the full potential of flexible adaptation. Table 1 categorizes quantitative methods applied to environmental policy design according to their ability to a) develop dynamic policy, b) incorporate obser vational data, and c) systematically consider future observations and strategy adjustments in current decision-making.

【a类方法的特点】Methods in category (a) can design a dynamic path for decisions over time. For example, ref. 13 used the dynamic programming (DP) method, a classical sequential decision-making framework, to estimate the optimal path of coastal protection based on current projections of the future climate. Ref. 14 employed heuristic algorithms, stochastically generating thousands of potential paths of the coastal protection and selecting better paths. These heuristic algorithms improve DP’s ability to handle the curse of dimensionality for multistep decision-making ( 14 ). 【然而，a方法存在局限不足:：决策路径多之后维数太高】Such methods, however, assume a static base of information and do not directly address a key advantage of flexible policy design: The capacity to learn and thus update and improve decisions as exogenous information is collected.

【b类方法的特点：动态观测】Methods in category (b) can design dynamic decision paths that adapt to new observations. In extreme cases, they account for the possibility that new information could shift scientific beliefs away from correct climate outcomes, a phenomenon known as negative learning ( 21 ). Specifically, Bayesian dynamic programming (BDP) methods ( 15 , 16 ) extend traditional DP by integrating new obser vations and projections into the estimation of optimal paths. 【b类方法的不足:贝叶斯动态规划，考虑当前状态，无法考虑未来】Although BDP can incorporate observations and learning into current decision-making, future learning and updating are not considered, and thus, lifetime risks to be addressed may be inac curately estimated. In other words, the potential for future decision adjustments can affect the optimality of current decisions

【另一类方法，决策树可以解决以上部分问题】Decision trees or real options methods which generate flexible plans by searching over scenario trees can overcome this limitation ( 7 , 17 , 18 ). 【然而，决策树存在方案指数增长】However, the real options approach involves an event tree with scenarios exponentially increasing with the total number of time steps in the policy pathway. Real options analysis is tractable only under a limited number of potential solutions and scenarios. Direct policy search (DPS) approaches ( 19 , 20 ) lower the compu tational cost by modeling the decision at each time step as a specified function of the observation, with the parameters of the function optimized through simulations. Consequently, the intricate stochas tic sequential decision is treated as a parameter optimization prob lem. Despite their computational efficiency, DPS approaches may still fail to deliver true optimality in adaptive climate decision-making.

【RL基本介绍】RL is an area of machine learning regarding how agents should act in changing environmental states to maximize their cumula tive rewards ( 22 ). RL approaches systematically incorporate observations, consider future outcomes and reactions, and sup port policy designs over a continuous range of future environ mental states. Also, various approximations (e.g., in characterizing states and rewards) can be made in RL for numerical efficiency.【RL的基本运用领域：chess, 水资源管理等】 RL has been successful in fields including chess playing ( 23 , 24 ), autonomous driving ( 25 ), and robotics control ( 26 ), and it has been employed to address sequential environmental decisions with large decision spaces, including electric power storage ( 27 ) and water resource management ( 28 ). 【RL在该领域没有得到应用，气候变化风险管理】However, it has not yet been used to address uncertainties in climate change risk man agement. We investigate the potential and performance of RL when applied to adaptive climate decision-making. Broadly speaking, we examine the value of systematic learning and updat ing in climate adaptation.

【我们采用RL作了案例】To illustrate, we apply the RL method to model adaptive strategies to address coastal flood risks. 【现实中对应的动作有哪些-对应RL中的action】Potential coastal adaptation strategies include “protection,” such as building structures such as seawalls along the coastline or dikes on higher lands; “accommodation,” such as retrofitting structures (encouraged through incentives and insur ance regulations); and “retreat,” or withdrawing from harm (which can be encouraged through subsidized “buyouts”) ( 29 ). 【未来情景变化的预估因素有两个：热带气旋和SLR】We account for projected changes in tropical cyclones (TCs), which may induce higher storm surges under climate change ( 29 – 33 ), and sea-level rise (SLR), which has been and will remain primarily responsible for increasing coastal flooding. 【然而，SLR的预估比较复杂】However, future SLR projections are characterized by large and deep uncertainties, associated par ticularly with human emissions and with ice-sheet physics, that currently impede modeling optimal risk mitigation strategies ( 29 , 34 – 38 ). 【利用RL来解决以下问题：动作的决策（适应，撤出、改造和筑堤）】We develop RL methods to identify optimal coastal risk mitigation strategies (including adaptive seawall construction and combined strategies involving withdrawing, retrofitting, and diking) for Manhattan, New York City, USA (NYC), that incorporate con tinuous SLR observations over the 21st century (Materials and Methods ).【RL算法带来的优越之处：解决了计算成本的问题】 RL efficiently handles the computational costs (which grow exponentially in traditional algorithms as the number of SLR scenarios and temporal resolution of decision updates increase) through state and reward approximation methods [Materials and Methods ; ( 39 )].

【我们聚焦于洪水风险管理领域，来评估其有效性】We focus on coastal flood risk management to evaluate the effectiveness of RL in the wider optimization frameworks for climate adaptation strategies. 【对比多种方法分析我们的RL方法优越之处】Comparison to DP, BDP, and DPS shows RL’s superiority in deriving flexible strategies that minimize cost and tail risk. 【RL方法在特殊情况也可以做得比较好】The RL framework also demon strates the greatest ability to minimize “regret” when the actual climate conditions deviate far from initial belief. The results highlight the importance of continuous learning and systematic adaptation in addressing the large uncertainty in climate pro jections and RL’s potential in modeling optimal climate adap tation strategies.

# Ste

## Social tipping dynamics for stabilizing Earth’s climate by 2050

【xx任务重要-防止危险的气候变化和毁灭性结果是决定性任务】Preventing dangerous climate change and its devastating con sequences is a defining task for humanity (1, 2). It is also anindispensable prerequisite for achieving sustainable development(3, 4).【引证-巴黎协定的要求-去碳化】 Limiting global warming to 1.5 °C as stipulated in theParis Climate Agreement (5) scientifically implies a complete netdecarbonization of the world’s energy and transport systems, in dustrial production, and land use by the middle of this century. Intheir “roadmap for rapid decarbonization,” 【去碳化如何具体做】Rockström et al. (6)highlight that rapid increase of the share of zero-carbon energywithin the global energy system would be needed to achieve thisobjective, likely alongside a considerable strengthening of terres trial carbon sinks. 【具体做的一种设想方法-目标展示分析】In one scenario, the zero-carbon share of theenergy system doubles every 5 to 7 y for the next several decades(6). Carbon emissions that are currently still on the rise at rates of0 to 2% per year, despite decades-long efforts in international cli mate negotiations, would thereby need to pivot to a rapid decline ofultimately 7% per year and more. 【这种方法中的目标是难以达到的，传统做不了】These emission reduction rateswould surpass by far even those experienced only during periodsof massive socioeconomic crisis in the 20th century, such as WorldWar II and the collapse of communism (Fig. 1).

【目前的关键问题是部署并实现它】Here, the historically decisive question is whether and how such rapid rates of deployment can be collectively achieved. 【目前的部署并实现有很多阻力】Current deployment rates of low-carbon energy sources are compatible with the required shift but when scaled up are expected to encounter considerable resistance due to the rigidities inherent in political and economic decision making (7, 8), as well as new technological demands (9, 10). 【虽然国家现在正在努力】Although an increasing number of countries have already introduced or are committed to introducing carbon pricing, the initiatives covered by carbon pricing included only 15% of global greenhouse gas emissions in 2017 (11) and have so far driven only marginal emission reductions (12).【然而，传统BAU的技术方案很难实现大量的减排目标】 It is increasingly recognized that business-as-usual technological progress and carbon pricing alone are not likely to lead to rapid and deep reductions in greenhouse gas emissions (13).

【观点引出，特殊情况下可以发现不同变化率】At the same time, there is evidence from various scientific fields demonstrating that rapid rates of change can be observed under certain critical conditions in natural (14–16), socioeconomic (17–20) and social-ecological systems (SESs) (21, 22). Increasing attention is being given to the concept of tipping dynamics as a nonlinear mechanism behind such disruptive system changes. Based on a review on social-ecological tipping points research, 【例子引出，milkoreit 提出概念】Milkoreit et al. (23) propose a common definition of social tipping points (STPs) as points “within an SES at which a small quanti tative change inevitably triggers a non-linear change in the social component of the SES, driven by self-reinforcing positive-feedback mechanisms, that inevitably and often irreversibly lead to a qual itatively different state of the social system.” There are historical examples of dynamic social spreading effects leading to a large self-amplification of small interventions: For example, the writings of one man, Martin Luther, injected through newly available printing technology into a public ready for such change, triggered the worldwide establishment of Protestant churches (24). 【气候领域的例子是通过关税、补贴和强制措施来刺激可再生能源】An exam ple in the field of climate policy is the introduction of tariffs, subsidies, and mandates to incentivize the growth of renewable energy production. This has led to a substantial system response in the form of mutually reinforcing market growth and expo nential technology cost improvement (25, 26)

【本文的工作是研究了去碳化的ste部分】In this paper, we examine a number of potential “social tipping elements” (STEs) for decarbonization (27, 28) that represent specific subdomains of the planetary social-economic system. Tipping of these subsystems could be triggered by “social tipping interventions” (STIs) that could contribute to rapid transition of the world system into a state of net zero anthropogenic green house gas emissions. 【ste确定的来源过程】The results reported in this study are based on an online expert survey, an expert workshop, and an extensive literature review (SI Appendix).

# 1+RL+ABM

# Stand2019-ays

## Deep reinforcement learning in WorldEarth system models to discover sustainable management strategies

I. INTRODUCTION

【在全球可持续发展的识别路径中需要考虑socio world和生物物理地球之间的互馈过程】Efforts invested in identifying pathways toward global sustainability need to account for critical feedback loops between the socioeconomic and sociocultural World and the biophysical Earth system.1,2 【发现这些路径需要考虑多个方面的领域知识，来达成safe just 空间】These pathways may require novel, yet undiscovered, multilevel policies, from the local to the global scale, for the governance of this coupled World-Earth system leading toward a safe and just operating space.3,4【为了这些目标，un指定了17个sdgs来实现决策一致】 Striving for a safe and just operating space, policymakers of the United Nations agreed on global political cooperation for a sustainable future at the resolution of the 17 Sustainable Development Goals (SDG)5 and the adoption of the Paris Agreement on Climate Change.6【safe just的空间理论基于生物物理地球边界（rockstorm）、raworth进行了拓展】 The safe and just operating space is based on a set of biophysical planetary boundaries (dened on dimensions such as climate change or biosphere integrity loss) as they are formulated by Rockström *et al.* in Refs. 3, 4, 7, and 8, extended by social foundations (e.g., poverty alleviation) by Raworth.9 If respected together, staying within these boundaries is seen as a prerequisite to ensuring sustainable human development. The eld of Earth system model ing develops computer models to show possible pathways toward a sustainable future. 【然而，决策里面的pb和sf轨迹难以定量描述】However, the identication and characterization of concrete trajectories within the planetary boundaries and above social foundations remains a problem requiring ongoing research e‑orts.10,1

【本文中的基本假设】In this paper, we consider this problem on a globally aggregated level assuming the following basic structure: An abstract single decision-maker interacts with a dynamical, in most cases, nonlinear environment to nd sustainable trajectories within certain boundaries. 【IAM是解决基本假设的问题的手段之一】The eld of Integrated Assessment Modeling (IAM) addresses this issue via optimizing a social welfare function in order to estimate the design of sustainable management strategies.12 IAM models integrate data and knowledge from established climate models.13,14 To identify pathways in IAM, numerical solvers such as GAMS15 are frequently used. 【然而，这些IAMS高度依赖目标函数来确定选择，存在缺陷】However, these IAM models are highly dependent on the choice of the target function of the optimization. In many cases, this choice may not be obvious and depends on the IAM developers.16 【OCT是解决类似问题的动态系统的一种方案】As another approach, optimal control theory (OCT) can be used to solve problems where dynamical systems are supposed to stay within certain constraints. In these systems, OCT tries to nd an optimal choice for some control variable by optimizing a specic objective function.17 Applied to Earth system models, the focus has been set on the design of climate regulators and their impact on climate modication.18,19 【oct方法中结合气候问题的子领域是viability theory】Viability theory (VT) as a subeld of OCT can be stated as an example. In this field, such problems of identifying trajectories are typically addressed by methods that rely on adiscretization of the state space, followed by the application of local linear approximations.20【然而，oct方法的不足是维数诅咒】 It is, however, not well applicable in systems with more than just a small number of variables due to the curse of dimensionality.2

【引出rl 是一种在world-earth系统中可行的方法】The use of reinforcement learning (RL)22 can also be considered as a possible approach for intelligent decision-making within World-Earth system models.23 It is designed for nding optimal policy strategies as well. However, in contrast to the previously presented approaches, RL does not detect solutions based on numerically solving an optimization problem, but by a dynamic search process via exploration and exploitation of past experiences, guided by a reward

function. 【传统rl方法不适用于该类问题】However, tabular methods, which are mainly used for classical RL solutions, cannot be straightforwardly applied to the systems of interest here, due to the continuous state spaces that we mostly nd in World-Earth system models.

The common point of all the presented methods outlined above is that they reach their limits as the complexity of the environments increases. However, deep reinforcement learning (DRL)24 algorithms have been shown to detect solutions in other highly complex environments surprisingly well.24,25 In this paper, we propose using DRL as a novel approach for Earth system analysis. Even though rst successful reinforcement learning experiments by using neural networks as nonlinear function approximators were reported already in 1995,26 the breakthrough of DRL was achieved only in 2013.24,25【drl是一种解决的好办法】 Since then, DRL algorithms have become increasingly popular in the eld of arti- cial intelligence.27,28 The key to the success of this approach lies in the combination of Q-learning,29 neural networks,30 and experience replay,31 which has been shown to learn policies up to a super human performance in a variety of di‑erent environments.24,25 Often DRL applications come up with unexpected and novel solutions.32,33 Many extensions have been proposed addressing both speed and eciency.34 【drl应用领域广泛】Due to its general applicability to various environments, DRL is used in a wide range of di‑erent elds, e.g., resources management in computer clusters,35 optimization of chemical reactions,36 playing abstract strategy games like chess and Go,32,33 autonomous driving,37 and, in particular, robotics.38–41

【基于drl的特点，我们提出了框架】Due to the wide applicability of DRL, we propose a framework that uses DRL as a tool that is both robust and easy to use at the same time to identify and classify trajectories in Earth system models e‑ectively. 【我们利用该框架在各种代表性模型中进行了实验】As a proof of concept, we use our DRL framework within various stylized World-Earth system models.2,42 【实验模型的特点是xx】These models are designed to investigate the coevolutionary dynamics of humans and nature in the Anthropocene. Some rst applications of reinforce

ment learning methods within resource use models have been carried out,43–45 【drl目前还没有用在地球系统模型的领域】but as far as we know, there are no approaches yet applying DRL to Earth system models.【该种方法的潜力即将释放】 We believe this approach will open so far unused possibilities to discover so far unknown management strategies that keep the Earth system within planetary boundaries, while, at the same time, respecting social foundations of the world’s societies. Recently, various ways of how to tackle problems related to anthropogenic climate change by using machine learning techniques have been outlined.46 Our work proposes a novel strand to this list.

II. METHODS

# C-切入点-ClimateRL-theo

## Can Reinforcement Learning support policy makers? A preliminary study with Integrated Assessment Models

**1 Introduction**

【气候是一个高维度系统切长时间依赖】Climate is a high dimensional dynamical system with strong inter-dependent components and long time dependencies, all of which interact to produce highly non-linear responses and behavior. 【气候和人类活动相关】Climate is also highly conditioned on human behavior – another greatly complex system – such that is now necessary to reason about climate change from a socio-climatic perspective [Moore et al., 2022]. 【为了找到解决方法，IAM被开发出来】To make progress towards achieving some kind of solution in the face of extreme consequences, policymakers and advisory groups employ *Integrated Assessment Models* (IAMs), state of the art models for climate change that combine knowledge about human development (such as economical theories) together with planetary sciences such as ecology and geophysics [Parson and Fisher-Vanden, 1997]. 【主流IAM模型采用大规模，计算代价高】Exploring and analysing the properties of the IAMs employed for large scale assessments [Pörtner et al., 2022] - to e.g. measure fidelity against the real world – is generally intractable from a computational perspective, which leads researchers to implement poor simplifying assumptions and decrease their effectiveness [Asefi-Najafabady et al., 2021]. 【简化的IAM也可以操作，类似于ode求解】Smaller IAM models aim to provide an alternative by employing fewer state variables and simpler sets of dynamics, making them amenable to mathematical probing and analysis [Kittel et al., 2017, Nitzbon et al., 2017]. Th literature commonly explores these models with ODE solvers; 【然而，最近证明可以用RL来解决研究系统和分析轨迹】however recently [Strnad et al., 2019] has shown that it is possible to employ them as environments in standard Reinforcement Learning (RL) [Sutton and Barto, 2020], and explore models using trained policies. 【RL扮演的潜能是提供复杂决策】These can be used to understand the system, and provide upwards insight towards improving more complex IAMs or even our understanding of climate change policies.

【我们基于前面工作来继续完善，测试算法，奖励函数和实验设计】We build on this work, testing more RL algorithms, reward functions, as well as different experimental setups. Among others, we show that (a) modern RL can learn effective policies with a variety of reward functions in this environment (b) that different agents and reward functions generate a significantly diverse set of solutions, thus exploring the IAM in different manners, and (c) that it is necessary to apply care when designing reward functions as they show different success rate in reaching the desirable state for different initialisation points, 【通过实验结果发现RL可以帮助更好理解挖掘应用模型】finally - (d) that RL helps us gain a deeper understanding of the properties and limitations of the applied models.

**2 AYS environment, RL reward functions and agent**

## Climate Change Policy Exploration using Reinforcement Learning

Introduction

1.1 Motivation

**【1.1.1 A Safe Space for Humanity】1.1.1 A Safe Space for Humanity**

【背景引出，人类活动影响自然环境】“It is unequivocal that human influence has warmed the atmosphere, ocean, and land. Wide spreadand rapid changes in the atmosphere, ocean, cryosphere, and biosphere have occurred.” according to the latest Intergovernmental Panel on Climate Change (IPCC) report [2].

【人类世-地质特点背景介绍】The relatively stable climate of the last 10-12 millennia, the Holocene, has seen the majority of human advancement: from the first agricultural societies to the start of the industrial revolution. The geological period before the Holocene, is known as the Pleistocene, colloquially known as the Ice Age [3]. Recently, due to the consequential effect of humans on the climate, scientists have started naming the era of human industrialisation the Anthropocene [4, 5]. This proposed geological era is characterized by the effects of humans on the environment and the climate.

【STE的引出，在人类世中ste很关键】Tipping Points refer to certain thresholds at which small perturbations can drastically change a system [6]. For the climate, different tipping elements can cause dramatic changes to the overall climate balance. According to scientists, the climate of the Holocene has a delicate balance [6, 7], one that should not be perturbed lest disastrous consequences for life on Earth, which would be difficult to counteract. It is understood that the climate is a part of a complex coupled socio-ecological system that is also adaptive [8]. 【人类行动会有很多影响，指定政策要考虑ste】So, any actions taken by humans can have many cascading and unpredictable effects [7]. Therefore, the coupled system has to be considered when constructing policy that relates to the climate [9].

【为了解决气候变化问题，PB的概念被提出来】To mitigate the effects of future climate change, scientists recommend staying within Planetary Boundaries [10, 11]. These boundaries define limits for measurable quantities, that are linked with significant tipping points, such as keeping warming below +2◦C above pre-industrial levels. 【PB越过的后果】Crossing planetary boundaries can provoke feedback loops (or “snowballing” effects) in the climate system that can lead to rapid and uncontrollable climate destabilisation [7]

**【1.1.2 Climate Change as a Control Problem】1.1.2 Climate Change as a Control Problem**

【ssp的引出，5种，预测5种代价很高】The IPCC uses SSP (Shared Socioeconomic Pathways) to make predictions on the possible pathways for the climate given different behaviours of humans. There are five pathways described, ranging from heavily reducing Green House Gas (GHG) emissions, to heavily increasing them. Each potential behaviour has consequences for the Earth and the life on it. Predicting these pathways requires models that have incredibly high uncertainties attached to them, since modelling all the complexities of the Earth is infeasible. The IPCC does not give exact probabilities for each scenario; instead, they predict using terms such as “very high probability”, “low probability”, “medium probability”. 【ssp来源，通过iam生成】These SSP are generated by Integrated Assessment Models (IAM). These models encompass a wide range of scientific knowledge in different fields such as climate, biology, and economics [12].

【气候系统的引出，特点是初值敏感】We know that the weather system and, consequently, the climate are dynamical systems where small changes in initial conditions can lead to startlingly different answers over time [13, 14]. Dynamical systems are deterministic. If the initial conditions are perfectly known, then we can predict outcomes accurately. Having fully accurate data and models is not technically possible, so meteorologists start multiple simulations at different initial conditions and then average the answers, also called ensembles models.【气候特点、ssp和IAM的联系，确定流程】 This is the approach also taken by the IPCC to generate SSP, they use information from multiple IAMs from the Integrated Assessment Model Consortium and average them out. In dynamical systems, this can yield an incredibly wide range of answers with enough time, which is why weather predictions are often inaccurate further in the future.

【气候的定义】Climate is defined as the “general weather conditions usually found in a particular place”1. The global climate is then the general weather patterns of the Earth. The global climate is easier to predict than the weather over long periods, since it is an average of the weather over long periods. 【然而，由于人类活动，气候被干扰】However, due to human influence, this system is being perturbed very fast [15]. Adding perturbations to a dynamical system makes the system even harder to predict. Perturbations can include a change in human carbon emissions, which has cascading and unpredictable effects. 【分析气候收到扰动后应该如何对应】Ideally, if we can predict accurately, we can control accurately. We would like to control the climate to negate its most disastrous effects for humans by staying within planetary boundaries. 【控制决定的重要性，决策实现可持续发展未来】Control would be done by policymakers and worldwide organisations by changing the actions of humans to enable a sustainable future.

【概念引出：保持动态系统在边界内叫做OCT】The task of keeping a dynamical system within certain boundaries is called Optimal Control Theory (OCT), often referred to more simply as control [16]. This sub-field of mathematics looks at methods that aim to optimise some objective function in a system by controlling the variables of the system. 【oct在具体气候问题定义中的描述】In the case of the climate, we would like to minimise climate destabilisation by staying within planetary boundaries

# C-切入点-Shuvo