Modeling Global Climate Negotiations and Cooperation in RICE-N

Tianyu Zhang, Andrew Williams, Yang Zhang, Prateek Gupta, Yoshua Bengio
MILA

Soham Phade, Sunil Srinivasa*, Stephan Zheng†
Salesforce Research

Abstract

Comprehensive global cooperation is essential to limit global temperature rise while maintaining sufficient long-term economic growth and reducing inequality. From a modeling point of view, achieving long-term global cooperation poses a complex game-theoretic problem with n competitive strategic agents that may negotiate, without a central authority. Hence, it is critical to design multi-lateral negotiations and agreements that best foster cooperation towards mitigating climate change while allowing all parties to meet their individual policy objectives. To address this interdisciplinary challenge, MILA and Salesforce Research are co-organizing a competition on designing contracts and negotiation protocols that foster global cooperation on mitigating climate change. Competitors will design and evaluate their solutions using an Integrated Assessment Model (climate-economic simulation) and a mix of machine learning tools, economic intuition, and domain knowledge. Here, we describe the technical background and implementation of RICE-N, the base IAM that is used for the design and evaluation of proposed solutions. In addition, we describe a reference multi-agent reinforcement learning implementation to train rational agents in RICE-N, both for a pure-CPU and a fast GPU-based version powered by WarpDrive.

1 Introduction

The latest IPCC report warns that it is “now or never" to stave off a climate disaster. The global climate system is approaching its tipping points. Ecosystems are drastically changing: the Amazon rainforest [1] is receding and polar ice sheets are melting [2, 3]. Other warning signs are also evident, such as the recent uptick in coastal flooding, forest fires and other extreme weather events. These developments are increasingly being attributed to climate change.

Climate change is a global phenomenon, and it affects all. In response, private and public financing have driven technological innovation (e.g., renewable energy transition) and community campaigns to drive systemic change. However, investments in mitigating climate change vary in size and type across nations and global regions, due to various social and economic factors. For example, developing nations may need to focus on the basic needs of their citizens first, while developed nations have more funding and opportunities to prepare for the adverse impacts of climate change. Hence, climate change presents a classic case of "tragedy of the commons", wherein individual agents pursuing their own self-interest may lead to a destructive outcome for everyone.

*Work done while at Salesforce.

†Email correspondence to: `stephan.zheng@salesforce.com`. The code is open-sourced at `https://github.com/mila-iqia/climate-cooperation-competition`. The most recent information and the latest version of this white paper can always be found at `https://mila-iqia.github.io/climate-cooperation-competition`. For detailed technical queries, join the Slack group: `https://join.slack.com/t/aieconomist/shared_invite/zt-g71ajic7-XaMygwNIup\protect\unhbox\voidb@x\protect\penalty\@M\{}CCzaR1TOwgA`. Disclaimer: This project is for research purposes only. We do not intend to make normative or value statements about different social welfare objectives or policies.

As a result, comprehensive global cooperation is essential to achieve the Paris Agreement's long-term goal of limiting global temperature rise to well below 2°C above pre-industrial levels [3]. At the same time, it is important to maintain sufficient long-term economic growth and reduce inequality. For example, public net-zero commitments can motivate domestic policy reforms or foreign investments and technology transfer to enable developing countries meet net-zero commitments in return for trade treaties. Such global cooperation may be implemented through climate clubs, for example, that may overcome barriers to take action on climate change.

From a modeling point of view, achieving long-term global cooperation poses a complex game-theoretic problem. This game can be modeled with n strategic agents, each agent representing a region or nation. Each agent is strategic and implements policies to achieve its own socio-economic and climate objectives, which may be at odds with other objectives and unknown to other agents. Agents interact through trade, diplomacy, foreign aid, and investments. Agents may be boundedly rational and its objectives may be hard to capture numerically. There is no central entity to enforce cooperation amongst agents. Instead, cooperation (if any) often originates from mutual negotiations and agreements. For these reasons, it is critical to design multilateral negotiations and agreements that best foster cooperation towards mitigating climate change while allowing all parties to meet their individual policy objectives.

To address this interdisciplinary challenge, MILA and Salesforce Research are co-organizing a competition on AI for mitigating climate change.

The objective of this competition is to design contracts and protocols that foster global cooperation on mitigating climate change. Competitors will design negotiation or other market mechanisms (please refer to section 5 in detail) and evaluate their solutions using a climate-economic simulation and a mix of machine learning tools, economic intuition, and domain knowledge.

In this white paper, we provide the technical background to participate in this competition. We describe the technical aspects of this simulation, game-theoretic description of global cooperation, example negotiation protocols, and ways to model the strategic behavior of agents. We also outline the evaluation process to assess submitted solutions and discuss relevant related work. Furthermore:

- In Section 4, we describe the environment dynamics of RICE-N, given the agent actions, and agent rewards.
- In Section 5, we describe how RICE-N can be augmented to facilitate negotiation between the regions.
- In Section 6, we describe how agents can be trained using RL.
- In Section 7, we describe the various metrics of interest such as how global temperatures behave over 100 years, total social welfare, and our evaluation criteria.
- We review several ethical considerations in Section 9.

2 Integrated Assessment Models

RICE-N is an instance of an Integrated Assessment Model (IAM). IAMs are commonly used in climate-change policy analysis to simulate future scenarios, based on a joint dynamics model of the climate and the economy. These underlying dynamics quantify the effect of economic activity and CO2 emissions on global temperatures and long-term economic development.

A pioneering example is the Dynamic Integrated model of Climate and Economy (DICE) [4]. DICE provides a simple dynamic representation of the scientific and economic links among various factors such as population growth, technological change, GHG emissions, carbon concentrations, climate change and damages. Although it provides a solid conceptual basis for IAMs, DICE only models a single global region. The Regional Integrated Climate-Economy (RICE) model, DICE's successor, generalizes this across multiple heterogeneous regions, accounting for differences in regional factors such as population growth or capital investments towards climate change mitigation. The RICE model was further extended to include international trades and tariffs [5].

IAMs are a great framework to model the positive impact of climate change mitigation and optimize climate policies. However, existing IAMs do not model strategic cooperation and competition between nations and economic regions, nor their negotiations and communication. These factors can significantly influence the short-term decisions of all stakeholders and their long-term effects. As such, interactions between agents are a critical but often overlooked aspect of climate-economic modeling.

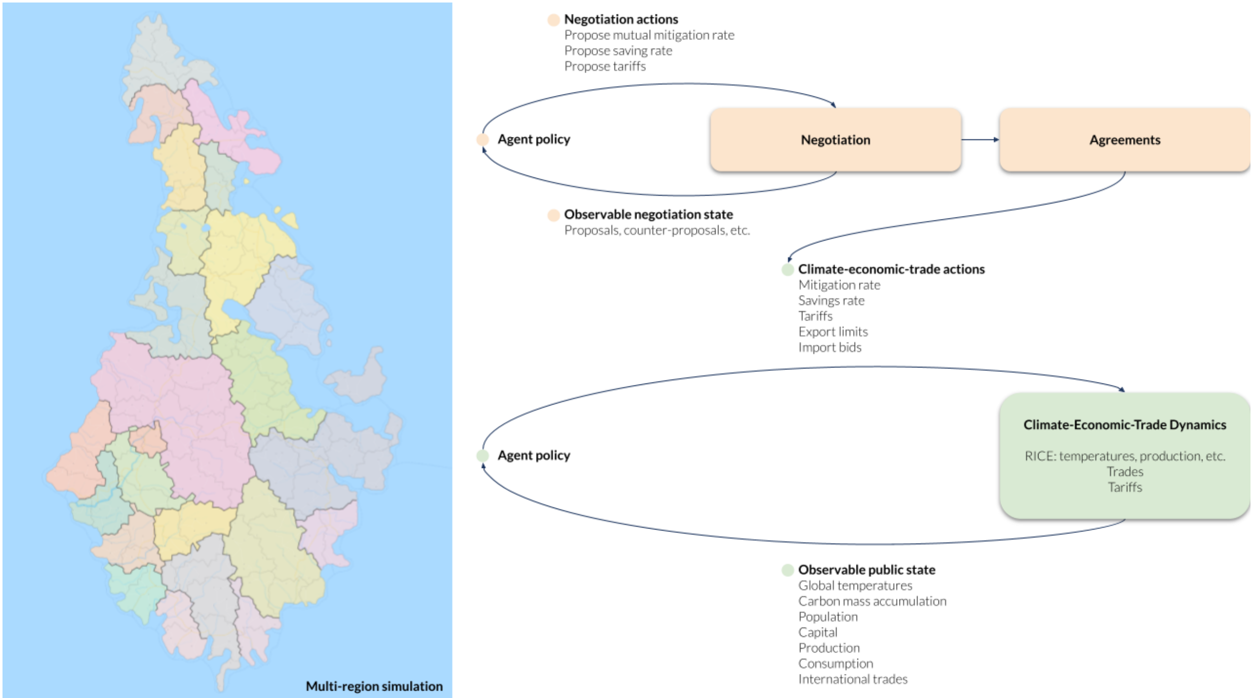


Figure 1: **Schematic overview of observables, actions, and components in RICE-N.** Each region (agent) uses a policy model to negotiate and make climate, economic, and trade decisions. For clarity, we show the flow of information and actions for a single agent only. The details of the negotiation protocol are up to competitors to implement. At a high-level, at each timestep, each policy first receives an observation of the negotiation and world state and decides how to negotiate. Negotiation may proceed for several iterations (at the same timestep). The outcome of all negotiations are agreements (or lack thereof), which may be between two or more agents. In particular, an agreement may influence the remaining actions that an agent can take in the climate and economic domains. For the same timestep, each agent then makes decisions with respect to the climate, economy, and trade.

3 The RICE-N Integrated Assessment Model

We introduce RICE-N, an IAM that further augments RICE with a framework for negotiation protocols (and further includes international trade and tariffs, as in [5]). Each agent (representing a fictitious region or nation) chooses its levels of capital investment, mitigation effort, and international trades and tariffs, for each time step (e.g., every 5 years). The simulation can be configured with an arbitrary number of agents and shares economic and social characteristics with the real world.

The structural parameters of the simulation have been fitted to real data. However, the public version of the simulation is not a one-to-one, identical representation of the real world. The individual agents do not represent real-world nations.

RICE-N is designed to be easily extendable and modular. In particular, RICE-N provides a framework to implement negotiation protocols as well as learning algorithms (e.g., reinforcement learning agents) to enable designing negotiation protocols and study their effects on global cooperation.

Base version of RICE-N. In RICE-N, there are n regions, where each region is modeled as an independent decision-making agent. These regions are fictitious. Regions interact with each other and the environment through their actions: setting a savings rate, mitigation rate, trades and tariffs, and negotiation actions.

The simulation has two main components: *negotiation* and *climate-economic activity*, see Figure 1. The activity component is the core component that simulates the physical actions of the agents and the resulting evolution of the environment. The negotiation component is an additional component that simulates communication between the regions that is necessary in forming agreements between the regions. In addition, agreements may adjust the available actions for each region during the activity stage.

Each simulation episode consists of H steps, each step representing Δ years (e.g, $\Delta = 5$). Thus, the simulation spans a period of $H \times \Delta$ simulation-years. At every step, the simulation goes through the negotiation stages and agreements are formed within the regions. The simulation then enters the activity stage where each region takes actions that comply with the agreements formed during the negotiation stages.

Climate-economic activity. The activity component of the simulation includes the following actions: the region's mitigation rates, and savings rates for capital investment, the tariffs imposed by this region on other regions, its maximum export limits, and import bids to other regions (see Table 2).

The activity component simulates the effect of the activity stage actions on the world climate and the economy. The *state of the world* is characterized by global variables such as the concentration of carbon dioxide in the Earth's atmosphere, and the average global temperature, as well as region specific variables such as population, capital, technology level, carbon intensity of economic activity, and balance of trade (see Table 1).

At every step, the simulator does the following (additional details provided in Section 4): The gross output production for each region is computed based on the state of the region, in particular, its capital investment, labor (or population), and technology factor. The net economic output is the gross output production reduced by climate damages from rising global temperature, and the cost of efforts towards mitigation by this region. The region consumes domestic goods equal to the quantity of the net economic output that is left after capital investment and export. It also consumes foreign goods from imports. The consumption utility for each region from consuming domestic and foreign goods is computed using the Armington elasticity assumption that has become standard in international computable general equilibrium models [6]. This gives the *reward* corresponding to each region in every step.

In economic terms, variables such as capital, balance of trade, carbon mass, and global temperature depend on the agents' actions and are called *endogenous* variables. On the other hand, variables such as population, technology level, carbon intensity of economic activity are called *exogenous*, i.e., their values do not depend on the agent actions. Note that the values of endogenous variables can vary across steps in a predetermined manner.

Negotiation. The negotiation component can include different stages of communication and the rules to form agreements between the regions. Note that the simulator can work with an empty negotiation component, and indeed these provide important baselines for performance evaluation. *The focus of this competition is to test different negotiation protocols that affect the actions during the activity stage and lead to socially better outcomes.*

In the base implementation of RICE-N, we allow the negotiated agreements to control the range of actions allowed for the regions. The effect of an agreement can then be encoded in the form of *action masks* that control the allowed action space during the activity stage. For example, an agreement could state that a given region should implement a minimum of 20% mitigation rate. Then the action mask only allows setting mitigation rates above that level for this region. Hence, it is important to ensure that the negotiation protocol provides each region a way to reject any agreement that restricts its actions. For example, it shouldn't happen that a certain region is forced to implement a mitigation rate that it did not agree upon. Alternatively, an implementation can include agent observations that allow regions to check if the agreements were followed and learn to punish the regions that did not comply with them. Such an implementation would be more realistic as compared to an implementation based on action masks because they do not assume any agreement enforcing entity.

Customization. The simulation code is structured so that custom negotiation protocols can be implemented as part of the negotiation component. An implementation may add agent observations and actions corresponding to the proposed negotiation protocol, and rules to generate the corresponding agreements. Notice that the agents can have access to these additional observations even during the activity stage. These observations are useful in letting the agents adopt policies in response to the negotiations and agreements. For example, they can learn to punish agents that do not comply with the agreements. Besides, by using action masking, the simulation can ensure that the agreements are followed (we can also have mix where some agreements are enforced using action masking and some others are non-enforceable). Thus, the simulation code provides complete freedom in implementing any negotiation protocol of choice provided the corresponding agreements are appropriately connected to the activity stage action through additional observations or action masks.

Consistency checks. On the other hand, for the purpose of the competition, we fix the actions in the activity stage and the corresponding climate and economic dynamics. Thus, certain parts of the simulation should not be modified by competitors, e.g., the core climate and economic parameters and equation, such as the ones which affect carbon emissions and productivity. Changing these components may change the problem and prevent fair comparisons across submitted solutions. As such, our evaluation protocol and implementation of RICE-N include consistency checks that test submitted solutions; each submission requiring submitting the full code of the (modified) simulation and agent

Variable	Type	Symbol	Description
Carbon Mass	Global, endogenous	M_t, $[M_t^{AT}, M_t^{UP}, M_t^{LO}]$	A three dimensional vector that indicates the average carbon accumulation in the atmosphere, upper oceans, and lower oceans.
Temperature	Global, endogenous	T_t, $[T_t^{AT}, T_t^{LO}]$	A two dimensional vector that indicates the average temperature of the atmosphere and the lower ocean.
Population	Regional, exogenous	$L_{i,t}$	Population and the labor in a region.
Technology	Regional, exogenous	$A_{i,t}$	Technology factor in the production function a region.
Capital	Regional, endogenous	$K_{i,t}$	Total capital accumulated by a region.
Carbon intensity of economic activity	Regional, exogenous	$\sigma_{i,t}$	A scalar coefficient that gives the emissions resulting from economic production.
Balance of trade	Regional, endogenous	$D_{i,t}$	Surplus or deficit from international trade activities.
Cost of mitigation efforts	Global, endogenous	$\theta_{1;i,t}$	An estimate of the cost of mitigation efforts.
Emission due to land use	Regional, exogenous	E_t^{Land}	Carbon emission for land use in a specific region.

Table 1: **World-state variables.** Global type variables correspond to the entire world, whereas regional type variables correspond to each region. Endogenous variables are those which are affected by the agent actions whereas exogenous variables are those that are predetermined and not affected by agent actions. Note that the values of endogenous variables can vary across steps in a predetermined manner. **Notation: indices are separated from subscripts referring to a name by semicolons (;). For instance, the parameter θ_1 varies in time t and by region i, which is denoted as $\theta_{1;i,t}$.**

Variable	Symbol	Description
Savings rate	$s_{i,t}$	The fraction of output production to be invested in capital.
Mitigation rate	$\mu_{i,t}$	The fraction of mitigation efforts by a region.
Import tariffs	$\tau_{i,j,t}$	The fraction of imports that are converted to tariff revenue.
Export limits	$p^x_{i,t}$	The fraction of domestic production that regions are willing to export.
Import bids	$b_{i,j,t}$	The amount of production each region is willing to import from other regions.

Table 2: **Agent-action variables.**

Variable	Symbol	Description
Initial population	$L_{0;i}$	The initial population for a specific region.
Population convergence target	$L_{a;i}$	The estimated convergence population for a specific region.
Population convergence rate	$l_{g;i}$	How fast the current population converges.
Initial capital	$K_{0;i}$	The initial capital for a specific region.
Initial carbon intensity	$\sigma_{0;i}$	The initial carbon intensity for a specific region.
Carbon intensity parameters	$g_{\sigma;i}$ and $\delta_{\sigma;i}$	The decay speed of the carbon intensity.
Initial technology factor	$A_{0;i}$	The initial carbon technology factor for a specific region.
Technology factor parameter	$g_{i,A}$ and $\delta_{i,A}$	The update pattern of the technology factor.
Initial land use emission	$E_{L0;i}$	The initial land use emission for a specific region.
Land use emission parameter	$\delta_{EL;i}$	The depreciation rate for the land use emission in a specific region.

Table 3: **Agent-specific constants.**

models. The activity stage is designed to follow the well established climate-economic dynamics in the RICE model augmented with trades and tariffs to include an essential influence in international negotiations. This way the simulation follows a simple model but is rich enough to implement interesting negotiation protocols.

Further, our implementation is set up to model each region as an RL agent that optimizes its discounted long term rewards. Although other agent policy models are possible, we require that the competitors clarify how their proposed agent policy models also aim to optimize for the same goal, namely, the personal discounted long term reward for each agent. This ensures that the agents behave greedily and any improvements in global metrics are strictly a result of the implemented negotiation protocols. See Section 6 for details.

Variable	Symbol	Description
Capital elasticity of production	γ	The contribution from capital and population to the economy.
Armington substitution parameter	λ	How substitutable consumption goods from different regions are.
Long term welfare discount rate	ρ	How much short-term welfare is weighted versus long-term welfare.
capital depreciation rate	Φ_K	The capital depreciation rate.
Backstop technology	p_b	Price of a backstop technology that can remove carbon dioxide from the atmosphere.
Backstop technology parameter	δ_{pb}	The decay speed of the cost of backstop technology.
Mitigation efficiency parameter	θ_2	The efficiency loss component of mitigation
Domestic share parameter	ψ^{dom}	The relative preference for domestic goods
Foreign share parameter	ψ^{for}	The relative preference for foreign goods

Table 4: **Global constants.**

4 The Activity Component: Climate, Economics, Trade, and Tariffs

The RICE-N simulation builds on the RICE model [7], a multi-region climate-economic simulation, by adding negotiation protocols, international trade, and support for strategic agents. RICE-N has n agents which each represent a region in the world, or a group of (fictitious) countries that are assumed to make decisions as a single entity. To maintain realism, we calibrate the structural parameters and establish a reasonable range of agent-specific parameters using data from the World Bank API [8].

4.1 Climate and Economic Dynamics

We now describe the RICE dynamics that govern the evolution of the world state from time t to $t+1$ for the different regions. The following equations determine the RICE dynamics (see also [9] for more details):

$$T_{t+1} = \Phi_T T_t + B_T \left(F_{2\times} \log_2 \left(\frac{M_t^{\text{AT}}}{M^{\text{AT},1750}} \right) + F_t^{\text{EX}} \right), \tag{1}$$

$$M_{t+1} = \Phi_M M_t + B_M \left(\sum_i \sigma_{i,t}(1-\mu_{i,t})Y_{i,t} + E_t^{\text{Land}} \right), \tag{2}$$

$$K_{i,t+1} = \Phi_{K,i} K_{i,t} + \Delta \left(1 - a_1 T_t^{\text{AT}} - a_2 \left(T_t^{\text{AT}}\right)^2\right) \left(1 - \theta_{1;i,t}\mu_{i,t}^{\theta_2}\right) Y_{i,t} s_{i,t}, \tag{3}$$

$$\theta_{1;i,t} = \frac{p_b}{1000 \cdot \theta_2} \left(1-\delta_{pb}\right)^{t-1} \cdot \sigma_{i,t}, \tag{4}$$

$$L_{i,t+1} = L_{i,t} \left(\frac{1 + L_{a;i}}{1 + L_{i,t}} \right)^{l_{g;i}}, \tag{5}$$

$$A_{i,t+1} = (e^{\eta} + g_{A;i} e^{-\delta_{A;i}\Delta(t-1)}) A_{i,t}, \tag{6}$$

$$Y_{i,t} = A_{i,t} K_{i,t}^{\gamma} L_{i,t}^{1-\gamma}, \tag{7}$$

$$\sigma_{i,t+1} = \sigma_{i,t} e^{-g_{\sigma;i}(1-\delta_{\sigma;i})^{\Delta(t-1)}\Delta}. \tag{8}$$

See Tables 1, 2, and 3 for the description of the variables and parameters in these equations. Variables without an agent index are global quantities. The values of some of the parameters are set following the data and procedure in [9] and [4], while others are calibrated using historical time series data (see the Appendix for more details).

At a high-level, Equations 1 and 2 capture climate dynamics (temperature and carbon mass), while Equations 3, 5, 6, and 7 capture economic dynamics. Finally, Equation 8 captures the carbon-intensity of production, providing a key link between the climate and economic sectors.

We now describe these equations in more detail.

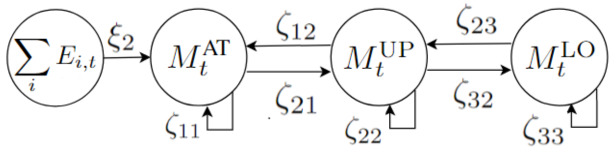


Figure 2: **The three-reservoir carbon mass model.**

4.1.1 Environmental Dynamics

Carbon mass. The total carbon mass in the climate system is given by:

$$M_{t+1} = \Phi_M M_t + B_M \sum_i E_{i,t}, \tag{2}$$

$$E_{i,t} = E_t^{\text{Land}} + \sigma_{i,t} A_{i,t}(1-\mu_{i,t})Y_{i,t}, \tag{9}$$

$$M_t \doteq \begin{bmatrix} M_t^{\text{AT}} & M_t^{\text{UP}} & M_t^{\text{LO}} \end{bmatrix}^\top \in \mathbb{R}^3, \tag{10}$$

$$\Phi_M \doteq \begin{bmatrix} \zeta_{11} & \zeta_{12} & 0 \\ \zeta_{21} & \zeta_{22} & \zeta_{23} \\ 0 & \zeta_{32} & \zeta_{33} \end{bmatrix}, \tag{11}$$

$$B_M \doteq \begin{bmatrix} \xi_2 \\ 0 \\ 0 \end{bmatrix}. \tag{12}$$

This describes a three-reservoir model of the global carbon cycle, in which M_{AT} describes the average mass of carbon in the atmosphere, M_{UP} is the average mass of carbon in the upper ocean, and M_{LO} the average mass of carbon in the deep or lower ocean, see Figure 2. Φ_M is the Markov transition matrix describing how carbon transfer between different reservoirs. B_M describes how the weight of carbon emission affects the carbon accumulation in the reservoirs.

Global temperature. Ultimately, increasing carbon mass leads to rising temperatures:

$$T_{t+1} = \Phi_T T_t + B_T F_t, \tag{13}$$

$$T_t \doteq \begin{bmatrix} T_t^{\text{AT}} & T_t^{\text{LO}} \end{bmatrix}^\top \in \mathbb{R}^2, \tag{14}$$

$$F_t = F_{2\times} \log_2 \left(\frac{M_t^{\text{AT}}}{M^{\text{AT,1750}}} \right), \tag{15}$$

$$\Phi_T \doteq \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}, \tag{16}$$

$$B_T \doteq \begin{bmatrix} \xi_1 \\ 0 \end{bmatrix}. \tag{17}$$

Similar to the carbon mass dynamic, there are two layers in the energy balance model, see Figure 3. T_{AT} is the combined average temperature in atmosphere, land surface, and upper ocean (simply referred to as the "atmospheric layer" hereafter). T_{LO} is the temperature in the lower ocean. Φ_T is the Markov transition matrix describing how heat transfers between different layers. B_T describes how carbon mass contributes to the temperature increases.

4.1.2 Economic Dynamics

Output production. The production in a region is given by the *total factor productivity (TFP)* [10] formula:

$$Y_{i,t} = A_{i,t} K_{i,t}^{\gamma} L_{i,t}^{1-\gamma}. \tag{7}$$

Production depends on three factors: total factor productivity ("technology") A_t, capital K_t, and labor L_t. This production function is common in the economic literature and used in the DICE/RICE models. The capital elasticity $\gamma \in [0,1]$ explains the different levels of contribution of capital and labor.

Population. The number of people in a region, denoted L_t grows as:

$$L_{i,t+1} = L_{i,t} \left(\frac{1 + L_{a;i}}{1 + L_{i,t}} \right)^{l_{g;i}}. \tag{5}$$

There are two parameters $L_{A;i}$ and $l_{g;i}$. $L_{a;i}$ represents the convergence population of region i and $l_{g;i}$ shows how fast the population $L_{i,t}$ converge to $L_{a;i}$. Please refer to the Appendix for a more detailed analysis and the calibration procedure.

Level of technology. The technology factor A_t describes how efficient production is, i.e., how many units of output a region achieves given fixed capital and labor:

$$A_{i,t+1} = (e^{\eta} + g_{A;i}e^{-\delta_{A;i}\Delta(t-1)})A_{i,t}. \tag{6}$$

Here, η represents the long-term growth of economics which is usually larger than 0, g_A represents the short-term part of economics growth, and δ_A represents the speed of decay of short-term growth factor. Δ is the time difference between steps. We use $\eta = 0.33\%$ as in [11].

Capital. The amount of capital evolves as:

$$\Phi_K \doteq (1-\delta_K)^{\Delta}, \tag{18}$$

$$K_{i,t+1} = \Phi_{K,i}K_{i,t} + \Delta\left(1 - a_1 T_t^{\text{AT}} - a_2\left(T_t^{\text{AT}}\right)^2\right)\left(1 - \theta_{1;i,t}\mu_{i,t}^{\theta_2}\right)Y_{i,t}s_{i,t}. \tag{3}$$

The evolution of the capital comes from two parts. The first part is capital inherited from the previous period with depreciation. In the second part, s_t is a control variable which represents the investment/savings rate (as a fraction of production). That is, as a base amount, the economy invests/saves a total of

$$Y_{i,t}s_{i,t}, \tag{19}$$

which yields new capital. This base amount is further modified by 2 multipliers: the damage function and mitigation/abatement costs, which are discussed below.

4.1.3 Climate-Economic Dynamics

Damage function. The climate damage function represents the economic damage due to climate change, e.g., increases in the atmosphere temperature T_t^{AT}. That is, in Equation 3, the fraction of new capital is modified by the damage function

$$1 - a_1 T_t^{\text{AT}} - a_2\left(T_t^{\text{AT}}\right)^2, \tag{20}$$

following [7]. That is, higher temperatures lead to less new capital. Similarly, $1-\theta_{1;i,t}\mu_t^{\theta_2}$ is the fraction of new capital after taking into account carbon emission mitigation. Mitigating carbon emissions more (higher μ_t) means (dirty) production needs to be lowered, hence yields less new capital.

Mitigation (abatement) cost. Following [9], for a mitigation rate $\mu_{i,t}$, the mitigation cost is

$$\theta_{1;i,t}\mu_{i,t}^{\theta_2}Y_{i,t}s_{i,t}, \tag{21}$$

where $\theta_{1;i,t}$ is given by Equation 4. This represents the loss in capital growth due to a fraction of production being used for mitigation.

Carbon intensity of economics activity. A critical part of the model is the interaction between the climate and economic parts. Specifically, the RICE model describes how production leads to carbon emissions:

$$E_t^{\text{Land}} = E_{L0}\cdot(1-\delta_{EL})^{t-1}, \tag{22}$$

$$E_{i,t} = E_t^{\text{Land}} + \sigma_{i,t}A_{i,t}(1-\mu_{i,t})Y_{i,t}, \tag{23}$$

$$\sigma_{i,t+1} = \sigma_{i,t}e^{-g_{\sigma;i}(1-\delta_{\sigma;i})^{\Delta(t-1)}\Delta}. \tag{8}$$

Here E_t^{land} is the carbon emission due to (changes in) land use, E_{L0} is the carbon emission in the base year, and δ_{EL} is the speed of decrease of changes in land use. The rates $0 < \delta_{EL} < 1$, $0 < \delta_{L0} < 1$ are free parameters. Due to a lack of data, E_t^{land} is set to be the same for each region.

$E_{i,t}$ is the total carbon emission, E_t^{land} is emission from natural sources, while $E_{i,t} - E_t^{\text{land}}$ is emission caused by economic activity. $\sigma_{i,t}A_{i,t}$ is the effective carbon intensity of economic activity: a higher technology factor lead to higher emissions, but can be modulated by lower σ (which can be thought of as the degree of "clean" production). $\mu_{idxi,t} \in [0,1]$ is a control variable called the abatement (ratio), which represents the proportion of the economics contributing to reducing carbon emission. Furthermore, we have 2 parameters g_σ and δ_σ that are fitted to data. g_σ is the rate of decrease in carbon emissions.

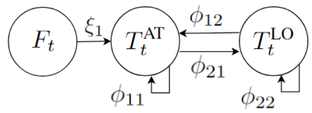


Figure 3: **The two-reservoir temperature model.**

4.2 Trade

We now describe the international trade dynamics and the resulting regional consumption and utilities. Regions trade by exporting their own consumption goods and importing other regions' consumption goods at a fixed unit price[3].

Agent actions. Each region i at time t must first specify a desired basket of consumption goods $\boldsymbol{b_{i,t}} = [b_{i,1,t}, ..., b_{i,k,t}]$ that they are willing to import from the other regions. These desired imports form a matrix of *bids* B_t such that the import bid by region i for goods from j at time t is $b_{i,j,t} \geq 0$, i.e., the amount of goods region i is willing to import from region j at time t is $b_{i,j,t}$.

Regions also set an upper bound $p^x_{i,t} \in [0,1]$ on the proportion of their own consumption goods that they are willing to export.

Trade constraints. To ensure that total imports and total exports match, three constraints are enforced on regions' trade flows.

1. For each region i, if the region's total desired imports from other regions exceed its own gross output, then the imports are scaled to sum up to the region's gross output. We enforce the constraint that $\sum_{i \neq j} b_{i,j,t} \leq Q_{i,t}$, which is to say that a region may not import more goods than its current gross output capacity. This constraint helps the agents avoid insurmountable debt, thereby stabilizing trade balances over the entire time period while also easing learning. If a region's desired imports exceed its production capacity, then its import bids are scaled down to size :
$$b_{i,j,t} \leftarrow b_{i,j,t} \min \left\{ 1, \frac{Q_{i,t}}{\sum_{i \neq j} b_{i,j,t}} \right\}. \tag{24}$$
2. Regions are allowed to carry a (positive or negative) trade balance $D_{i,t}$. At the start of each new time step, each region's trade balance, positive or negative, accumulates interest at a fixed rate of 10%. Based on this balance, a region's debt-to-initial-capital ratio is determined and the imports are scaled according to this ratio:
$$d_{i,t} = 10 \frac{D_{i,t}}{K_0}, \tag{25}$$
$$b_{i,j,t} \leftarrow b_{i,j,t}(1 + d_{i,t}). \tag{26}$$
3. If other regions' total desired imports from region i exceed region i's upper bound on exports $x^{\max}_{i,t}$, then the bids for goods from region i are scaled proportionally to $x^{\max}_{i,t}$. Otherwise, each region receives its full import bid from region i. In other words, region i cannot export more goods at time t than it could consume at time t, so other regions will import less from region i.
$$x^{\max}_{i,t} = \min(p^x_{i,t} Q_{i,t}, Q_{i,t} - I_{i,t}), \tag{27}$$
$$x_{i,j,t} = b_{i,j,t} \min \left\{ 1, \frac{\min(x^{\max}_{i,t})}{\sum_{j \neq i} b_{i,j,t}} \right\}. \tag{28}$$

After all constraints have been applied, the trade balance for the next period is calculated:
$$D_{i,t+1} = D_{i,t} + \Delta \left(\sum_{j \neq i} x_{j,i,t} - \sum_{j \neq i} x_{i,j,t} \right). \tag{29}$$

[3] More generally, prices should be dynamic, but the current implementation does not support this.

Algorithm 1 Activity Component (implemented by `Climate_and_economy_simulation_step()`) Note that we only list input state variables and omit model parameters.

```
Require: exogenous emissions, land emissions, intensity, production factor, labor, capital,
  previous global temperature, previous government balance
Require: actions : mitigation rates, saving rates, tariffs, export rate limit, desired imports
  for each region do
    mitigation cost ← f(intensity)                                          ▷ Equation 4
    damages ← f(previous global temperature)                                ▷ Equation 20
    abatement cost ← f(mitigation rate, mitigation cost)                    ▷ Equation 21

    production ← f(production factor, capital, labor)                       ▷ Equation 7
    gross output ← f(damages, abatement cost, production)                   ▷ Equation 7
    government balance ← f(interest rate, previous government balance)
    investment ← f(saving rate, gross output)                               ▷ Equation 19

    scaled imports ← f(gross output, desired imports)                       ▷ Equation 24
    debt ratio ← f(previous government balance)                             ▷ Equation 25
    scaled imports ← f(scaled imports, debt ratio)                          ▷ Equation 26
  end for
  for each region do
    max potential exports ← f(gross output, investment, export rate limit)  ▷ Equation 27
    Scaled imports ← f(scaled imports, max potential exports)               ▷ Equation 28
  end for
  for each region do
    tariff-ed imports, tariff revenue ← f(scaled imports, tariffs)          ▷ Equation 30
    domestic consumption ← f(savings, gross output, scaled imports)         ▷ Equation 31
    aggregate consumption ← f(domestic consumption, tariff-ed imports)      ▷ Equation 32
    utility ← f(labor, aggregate consumption)                               ▷ Equation 34
    government balance ← f(imports, exports)                                ▷ Equation 29
  end for
  temperature ← f(previous temperature, previous carbon mass, exogenous emissions)
  carbon mass ← f(previous carbon mass, intensity, mitigation rate, production, land emissions)
  for each region do
    capital ← f(capital, investment)                                        ▷ Equation 3
    labor ← f(labor)                                                        ▷ Equation 5
    production factor ← f(capital)                                          ▷ Equation 6
    carbon intensity ← f(carbon intensity)                                  ▷ Equation 8
  end for
```

Tariffs. Regions can also choose to impose import tariffs on other regions. We denote an import tariff imposed by region i on a region j by $\tau_{i,j,t} \in [0, 1]$. If region i imposes an import tariff $\tau_{i,j,t} \in [0, 1]$ on region j, region i consumes

$$C_{i,j,t} = x_{i,j,t}(1 - \tau_{i,j,t}), \tag{30}$$

and $\tau_{i,j,t}x_{i,j,t}$ is added to a reserve fund specific to that region.

Consumption. Consumption of domestic goods $C_{i,i,t}$ is determined according to gross output, the savings rate and exports:

$$C_{i,i,t} = (1 - s_{i,t})Q_{i,t} - \sum_{j \neq i} x_{j,i,t}. \tag{31}$$

The aggregated consumption $C_{i,t}$ at time t for region i is given by the Armington elasticity model [5]) as follows:

$$C_{i,t} = \left(\psi^{dom}(C_{i,i,t})^{\lambda} + \sum_{j \neq i} \psi^{for}(C_{i,j,t})^{\lambda}\right)^{\frac{1}{\lambda}}. \tag{32}$$

5 The Negotiation Component

A wide range of protocols can be implemented as part of the negotiation component. One of the goal of this competition is to design a negotiation mechanism between different regions so that agents may not only maximize their own utility but also care more about the collective goal — carbon emission mitigation.

We encourage but do not force the proposed negotiation mechanisms to be

- incentive-compatible: A mechanism is called incentive-compatible (IC) if every participant can achieve the best outcome to themselves just by acting according to their true preferences [12]
- decentralized: If no agents are able to make rules from its single side, the mechanism is decentralized
- strategy-proof: A mechanism is called strategy-proof given no information about what the others do, the agents fare best or at least not worse by being truthful. [13]

We explain the general structure through several naive examples.

5.1 No Negotiations

Let us begin with the most simple negotiation protocol, the one with no negotiation. In this case the simulation runs through the activity component at each step. Each region chooses its actions from the entire feasible range of actions without any restriction since no agreements are formed. Each region aims to optimize for its individual rewards and thus this is the classic case of tragedy of the commons.

5.2 Unilateral Contracts

The negotiation component could include various constraints that do not require negotiation. For example, if two regions a and b enact mitigation rates μ_a and μ_b such that $\mu_a > \mu_b$, then the negotiation component could enforce that region a imposes a tariff $\tau_a = \alpha(\mu_a - \mu_b)$ on region b where α is a "mitigation correction" coefficient. This is a simple mechanism through which a region a could incentivize region b to mitigate more.

5.3 Bilateral Negotiations on Mitigation Rates

The negotiation component can be composed of multiple stages of observations and actions that ultimately lead to agreements within the regions. Consider the following protocol: For each ordered pair of regions (i, j), region i makes a proposal $(\hat{\mu}_i, \hat{\mu}_j)$ and region j decides whether to accept the proposal or not. The proposal above means that, if accepted by region j, then an agreement is formed between region i and region j that says, region i will choose a mitigation rate μ_i at least as large as $\hat{\mu}_i$ and region j will choose a mitigation rate μ_j at least as large as $\hat{\mu}_j$ as their inputs to the activity component.

Thus, during the negotiation component, there will be $n(n-1)$ total proposals and decisions regarding their acceptance. After the negotiation component, each region i will select its mitigation rate μ_i that is greater than or equal to all of the mitigation rates $\hat{\mu}_i$ in the accepted proposals that region i was a part of.

This negotiation component can be implemented in two stages:

- *Proposal stage*: At this stage each region i makes a proposal to every other region j.
  Observations: state observations for agent i.
  Actions: Proposals $(\hat{\mu}_i, \hat{\mu}_j)$ for every other region j.
- *Evaluation stage*: At this stage each region observes the proposals made to it in the preceding stage and takes an action of accepting or rejecting each of the proposals.
  Observations: state observations for agent i, incoming and outgoing proposals for agent i.
  Actions: Accept or reject each received proposal.

At the end of the evaluation stage the minimum required mitigation rate is computed for each region and the corresponding action mask is set. This ensures that the regions choose actions for the activity component that comply with the agreements formed during the negotiation component. Instead of using action masks, the implementation can allow the agents to take any action but include these actions as observations for the other agents so that they can learn to punish any agreement violations. For example, the regions can increase the tariff rates on regions that violate agreements.

5.4 Multilateral Negotiations

Multilateral negotiations imply communication amongst several regions simultaneously. Although more technically challenging to implement, such negotiation protocols offer more possibilities and added realism. A key goal of this competition is to explore and implement effective negotiation protocols. We now discuss a recent proposal as an example.

Climate clubs. Climate clubs [7] are one example of a multi-region agreement. Each region that is a part of a club agrees to 1) enact a minimum mitigation rate and 2) enforce a minimum import tariff on regions that are not part of the club. A high minimum import tariff acts as an incentive to join the climate club by applying a mitigation rate at least as high as the club's minimum.

The negotiation component would specify how clubs are formed, e.g., by randomly selecting one region to make a proposal (μ_{min}, τ_{min}) to which all other regions can agree (thereby joining the climate club) or refuse.

This negotiation component can be augmented in many ways, e.g., by repeating it with the standalone regions until all regions form a climate club, by implementing an iterative proposal-evaluation framework, or by specifying a more sophisticated mechanism for climate club proposals than random selection.

Asymmetric negotiation protocols. Certain negotiation protocols such as climate clubs rely on the uniform treatment of trading partners to simplify their theoretical analysis [7]. However, such limits do not apply to agent-based simulations [14], and so the negotiation actions of regions in the simulator are not required to be applied uniformly across trading partners.

In fact, climate justice, on which effective international cooperation rests [15], often requires that countries be treated differently based on characteristics such as historical cumulative emissions or exposure to potential climate damages [16]. For example, article 4.4 of the Paris Agreement states that developed regions, which have historically contributed emissions in a disproportionate manner relative to developing countries, should continue taking the lead when it comes to mitigation efforts [17]. However, there is no agreed-on way to label regions as "developed" and "developing" in simulations such as RICE-N. Rather, we invite the community to consider how the characteristics of the fictitious regions in RICE-N relate to the climate debate in the real world.

Moreover, the "fairness" of climate change legislation can be established through many different methodologies which treat regions differently [18]. We encourage participants to explore important opportunities for the innovation and study of negotiation protocols that implement climate justice (as defined by the participants). Although building the simulator requires making design choices, we emphasize that we do not seek to make normative statements about "fairness" and "justice" in this work.

6 Modeling and Optimizing the Behavior of Agents

Each region can be modeled as a (boundedly) rational strategic agent. The objective of a rational agent i is to optimize its policy $\pi\left(a_t|o_t\right)$ to maximize the long-term aggregate γ-discounted utility for the corresponding region:

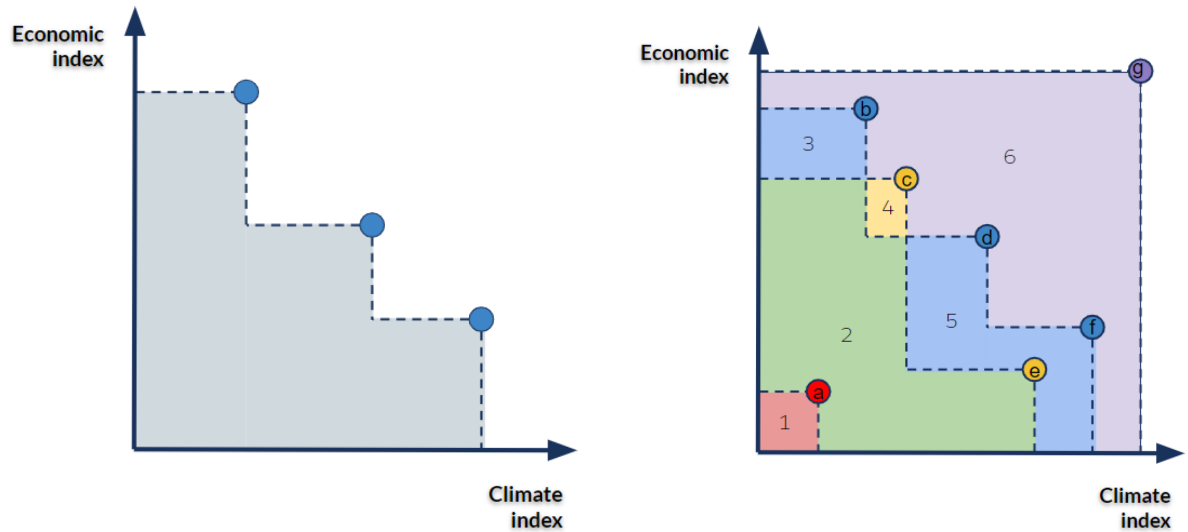
$$\max_{\pi_i} \mathbb{E}_{\pi_1,\ldots,\pi_n}\left[\sum_{t=0}^{T}\gamma^t r_{i,t}\right], \tag{33}$$

$$r_{i,t} = U_{i,t} = \frac{1}{1-\alpha} L_{i,t}\left(\frac{C_{i,t}}{L_{i,t}}\right)^{1-\alpha}, \tag{34}$$

$$C_{i,t} = \left(\psi^{\text{dom}}(C_{i,i,t})^{\lambda} + \sum_{j\neq i}\psi^{\text{for}}(C_{i,j,t})^{\lambda}\right)^{\frac{1}{\lambda}}, \tag{32}$$

$$C_{i,j,t} = x_{i,j,t}(1-\tau_{i,j,t}) \quad \forall j \neq i. \tag{30}$$

where $L_{i,t}$ is given by equation (5). Here, the actions for each agent include setting the mitigation rate $\mu_{i,t}$ and savings rate $s_{i,t}$, as well as trading and negotiation actions including import tariffs $\tau_{i,j,t}$, export limits $p^x_{i,t}$ and import bids $b_{i,j,t}$



(i) (ii)

Figure 4: **A visualization of how teams are ranked using the hypervolume indicator [20].** (i) A single team's set of submissions is represented by a set of points of the same color. The hypervolume indicator is the area of a set of non-dominated policies [21] in the 2-dimensional $(\mathcal{C}, \mathcal{E})$-space, as represented by the shaded area. The reference points (ii) Multiple sets of solutions proposed by various teams. Teams are ranked by the value of their hypervolume indicator: a higher value is a better score. A first team submits point (a), so their set of solutions contains one point. The hypervolume indicator of this solution corresponds to area 1. A second team submits points (c) and (e). Their score corresponds to the sum of areas 2 and 4. A third team submits points (b), (d) and (f), and so their score is the sum of areas 2, 3 and 5. A final team submits a single point (g). The hypervolume indicator of this point corresponds to the sum of all areas 1 to 6. As such, the team that submitted point (g) is ranked highest, the team that submitted points (b), (d) and (f) is second, the team that submitted points (c) and (e) is third, and the team that submitted point (a) is last.

such that $\pi_i = (\mu_{i,t}, s_{i,t}, \tau_{i,j,t}, p^x_{i,t}, b_{i,j,t})_{t=0}^T$. In the first iteration of this competition, we focus on the effect of various negotiation protocols and their effectiveness assuming regions are rational agents.

Given a simulation and negotiation protocol, finding the optimal agent behavior can be naturally done using reinforcement learning (RL). However, we encourage you to try different behavioral optimization algorithms, provided the implementation uses the same observations and actions, and aim to optimize for the same individual rewards.

Reinforcement Learning Agents. We provide an RL implementation to train rational agents for each region using A2C or PPO [19]. We encourage you to expand and extend this implementation.

The base implementation models each RL agent using a neural network policy that shares weights across agents, but uses region-specific inputs. The architecture of the network can be adjusted, e.g., the number of layers and the dimension of each layer. Agent policies use separate heads for each action.

To distinguish between agents, the policy model's input contains agent-specific features such as their population, capital, technology factor, damage function, as well as a one-hot representation of the region's index. Agents also receive information related to negotiation, e.g., the latest proposals made to this region and by this region, or the minimum mitigation rate agreed upon by this region in bilateral negotiations. In addition, each agent observes the public state of the world (e.g., climate conditions) and the current stage in the simulation.

The output of these regions is a combination of all the different actions for this agent across all stages, e.g., proposals for other regions, decisions on proposals made by other regions, and mitigation and savings rates. Note that not all observations and actions are relevant to each agent in every stage. Hence, the implementation only fetches the actions pertaining to that stage.

7 Evaluation Protocol

Structure of a submission. Submissions consist of the full simulation code, agent behavioral models, and other components that may be needed to generate full simulation roll-outs.

CPU vs GPU simulations. We provide both CPU and GPU versions of the base RICE-N, but evaluate all submissions using the CPU version to reduce the cost of evaluation. The CPU and GPU implementations of the base version of RICE-N yield consistent results. The GPU version of the simulation is meant to accelerate training. If participants use the GPU version, they should also provide a CPU version and ensure that any models trained on the GPU version and any customization yield consistent results on a CPU.

Evaluation metrics. For each submission, we evaluate the mean simulated global climate and economic outcomes, quantified through both climate and economic indices, across five random seeds. The climate index $\mathcal{C}$ measures the temperature change over the course of 100 years, *compared to two extremal policies that do not mitigate at all or use a 100% mitigation rate*. The economic index $\mathcal{E}$ measures the increase in global productivity (i.e., total GDP), similarly compared to two extremal policies. We compare against extremal policies to get normalized, dimensionless metrics along both climate and economic dimensions, which allows for fair comparison and computation of the hypervolume indicator. In particular, a basic intuition is that these two extremal policies represent two extremal points in the trade-off between the climate and the economy. However, these extremal policies may not be Pareto-optimal, may not be optimal under every negotiation protocol, and may not be individual best-response strategies.

Each submission thus yields a point in the 2-dimensional $(\mathcal{C}, \mathcal{E})$-space. The full set of submissions, collected across all competitors, yields a point cloud in this space. Within this space, the Pareto frontier is the set of all solutions for which neither index can be increased without decreasing the other index. We say that a solution a *Pareto-dominates* another solution b if at least one index is strictly higher and the other index is at least as high, i.e., 1) $\mathcal{C}_a > \mathcal{C}_b$ and $\mathcal{E}_a \geq \mathcal{E}_b$, or 2) $\mathcal{C}_a \geq \mathcal{C}_b$ and $\mathcal{E}_a > \mathcal{E}_b$.

Each team's score is defined as *the hypervolume indicator of the set of solutions which they provide*. The hypervolume of a set of solutions is the area of a set of non-dominated policies in the objective space with respect to a predetermined reference point [21]. In the case of our simulations, the reference point is determined by the extremal policies. Aside from providing a scalar value to compare submissions in a multi-objective setting, this score definition has several attractive properties: 1) The hypervolume indicator encourages submissions to lie as close as possible to the Pareto frontier, 2) it is strictly monotonic with respect to Pareto dominance such that a set of solutions that Pareto-dominates dominates another set must have a bigger hypervolume, and 3) it incentivizes diversification (along the Pareto frontier), i.e., teams are incentivized to submit solutions that make different climate-economic trade-offs to increase the hypervolume of their set. See Figure 4 for more details.

Randomness. The behavior of the climate-economic parts of the simulation is deterministic. However, the agents' behavior may be stochastic. Hence, we produce 10 simulation roll-outs under 3 random seeds, and determine the mean and standard deviation of the outcomes for evaluation purposes.

8 Discussion

We presented the RICE-N simulation and a conceptual framework to include strategic negotiation into climate-economic modeling. To our knowledge, this is the first version of a simulation that integrates negotiation with climate and economic dynamics. Given the modular structure of our implementation, RICE-N can be naturally extended to include more complex climate, economic, or strategic features in the future. As such, we believe this framework is a compelling tool to study AI for social good and invite the community to build on this work.

The RICE climate and economic dynamics are relatively simple, but are a minimal implementation of an IAM that supports studying the strategic aspects of decentralized decision-making, negotiation, and their impact on the climate. Both the climate and economic components could be made more sophisticated, as implemented by a large number of existing IAMs. For example, one might add rich spatial climate dynamics, or model more parts of the real economy, e.g., international companies. The game-theoretic features could also be enhanced, e.g., by allowing for forms of communication, more policy levers, more possible interactions (e.g., technology transfer), and others. However, simulating more complex dynamics also incurs significant computational cost. As such, we hope that future work will extend the RICE-N model towards using AI for climate change mitigation and social good.

Acknowledgements

We thank Vincent Conitzer, Richard Socher, Nicholas Kumbleben, and Eyck Freymann for valuable discussions and feedback. We thank Anna Bethke for the ethical review.

Contributions

- YB, SZ, PG conceived and directed the project;
- All the authors developed the theoretical framework;
- TZ, AW, SP, SS developed the economic simulator, implemented the reinforcement learning platform;
- SS developed the evaluation pipeline.
- SS developed the GPU version of the simulation.
- TZ scraped data;
- TZ, SP performed calibration;
- TZ, AW, SP, SS performed experiments;
- TZ, SP, AW developed the tutorial notebook;
- PG, YZ planned and advised the work;
- All the authors drafted the manuscript;
- All authors discussed and commented on the manuscript.
- All authors reviewed the code.
- PG, SZ developed the project website.

9 Ethical Considerations

While the intention of this paper and the corresponding challenge is to stimulate innovative solutions to climate change, there are some unintended consequences that we would like to acknowledge and address here. These include the carbon footprint of running the simulation itself, the economic disparities that can exist with climate negotiations, and the potential extensibility of this simulation to the real world.

Carbon emissions from simulations. First, it is important to acknowledge that running climate change simulations in RICE-N will inevitably release carbon emissions into our atmosphere. While the computational requirements of these simulations are much smaller than training large language models, they still exist. To mitigate this harm, we are encouraging participating teams to consider their energy use during experimentation, offsetting their own carbon emissions if possible.

Economic and climate disparities. Second, as the World Bank states "Climate change is deeply intertwined with global patterns of inequality" and yet "the most vulnerable are often also disproportionately impacted by measures to address climate change" [22]. While it is important to determine ways that to mitigate climate change it is equally important to ensure that vulnerable populations are not negatively impacted by climate change measures.

Technical aspects and limitations. Last but not least, it is important to note that the predicted climate and economic predictions made in RICE-N may differ in a real-world setting due to externalities beyond the boundaries of the simulation. A fictional world is utilized in this competition to further illustrate the potential gap between simulation and reality, but the uncertainty of the results should be fully understood, especially before implementing any policies recommended by RICE-N.

References

[1] Thomas E Lovejoy and Carlos Nobre. Amazon tipping point, 2018.

[2] Niklas Boers and Martin Rypdal. Critical slowing down suggests that the western greenland ice sheet is close to a tipping point. *Proceedings of the National Academy of Sciences*, 118(21), 2021.

[3] Robert M DeConto, David Pollard, Richard B Alley, Isabella Velicogna, Edward Gasson, Natalya Gomez, Shaina Sadai, Alan Condron, Daniel M Gilford, Erica L Ashe, et al. The paris climate agreement and future sea-level rise from antarctica. *Nature*, 593(7857):83–89, 2021.

[4] William D Nordhaus. A review of the stern review on the economics of climate change. *Journal of economic literature*, 45(3):686–702, 2007.

[5] Kai Lessmann, Robert Marschinski, and Ottmar Edenhofer. The effects of tariffs on coalition formation in a dynamic global warming game. *Economic Modelling*, 26(3):641–649, 2009.

[6] Paul S Armington. A theory of demand for products distinguished by place of production. *Staff Papers*, 16(1):159–178, 1969.

[7] William Nordhaus. Climate clubs: Overcoming free-riding in international climate policy. *American Economic Review*, 105(4):1339–70, 2015.

[8] WorldBank. The world bank documents and report api, 2022.

[9] Christopher M Kellett, Steven R Weller, Timm Faulwasser, Lars Grüne, and Willi Semmler. Feedback, dynamics, and optimal control in climate economics. *Annual Reviews in Control*, 47:7–20, 2019.

[10] Diego Comin. *total factor productivity*, pages 260–263. Palgrave Macmillan UK, London, 2010.

[11] William Nordhaus. Evolution of modeling of the economics of global warming: changes in the DICE model, 1992–2017. *Climatic Change*, 148(4):623–640, June 2018.

[12] Wikipedia contributors. Incentive compatibility — Wikipedia, the free encyclopedia, 2021. [Online; accessed 9-July-2022].

[13]

[14] J Doyne Farmer, Cameron Hepburn, Penny Mealy, and Alexander Teytelboym. A third wave in the economics of climate change. *Environmental and Resource Economics*, 62(2):329–357, 2015.

[15] Sonja Klinsky, Timmons Roberts, Saleemul Huq, Chukwumerije Okereke, Peter Newell, Peter Dauvergne, Karen O'Brien, Heike Schroeder, Petra Tschakert, Jennifer Clapp, et al. Why equity is fundamental in climate change policy research. *Global Environmental Change*, 44:170–173, 2017.

[16] United Nations. General Assembly et al. Report of the united nations conference on environment and development. In *A/CONF*, volume 151, page 26, Rio de Janeiro, 1992.

[17] Carl-Friedrich Schleussner, Joeri Rogelj, Michiel Schaeffer, Tabea Lissner, Rachel Licker, Erich M Fischer, Reto Knutti, Anders Levermann, Katja Frieler, and William Hare. Science and policy characteristics of the paris agreement temperature goal. *Nature Climate Change*, 6(9):827–835, 2016.

[18] Kate Dooley, Christian Holz, Sivan Kartha, Sonja Klinsky, J Timmons Roberts, Henry Shue, Harald Winkler, Tom Athanasiou, Simon Caney, Elizabeth Cripps, et al. Ethical choices behind quantifications of fair contributions under the paris agreement. *Nature Climate Change*, 11(4):300–305, 2021.

[19] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

[20] Eckart Zitzler, Lothar Thiele, Marco Laumanns, Carlos M Fonseca, and Viviane Grunert Da Fonseca. Performance assessment of multiobjective optimizers: An analysis and review. *IEEE Transactions on evolutionary computation*, 7(2):117–132, 2003.

[21] Kristof Van Moffaert and Ann Nowé. Multi-objective reinforcement learning using sets of pareto dominating policies. *The Journal of Machine Learning Research*, 15(1):3483–3512, 2014.

[22] WorldBank. Social dimensions of climate change, 2022.

[23] Lars Buitinck, Gilles Louppe, Mathieu Blondel, Fabian Pedregosa, Andreas Mueller, Olivier Grisel, Vlad Niculae, Peter Prettenhofer, Alexandre Gramfort, Jaques Grobler, Robert Layton, Jake VanderPlas, Arnaud Joly, Brian Holt, and Gaël Varoquaux. API design for machine learning software: experiences from the scikit-learn project, 2013.

[24] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.

A Creating a 27-Region Simulation

We feature $n = 27$ fictitious regions in our public simulation. These are inspired by merging and splitting real-world countries, but are not exactly the same as the real-world counterparts.

We used real data from the World Bank API [8], e.g., GDP, capital stock, population, and CO2-quivalent (CO2eq) emissions. Furthermore, the World Bank groups countries into regions, including Sub-Saharan Africa, South Asia, North America, the Middle East and North Africa, Latin America and the Caribbean, Europe and Central Asia, East Asia and Pacific. In each region, the different countries (or sub-regions) are classified into 4 income groups: high income, upper middle income, lower middle income, and low income.

Merging regions. We assume the GDP, capital stock, and population for the regions are additive. We also assume the gross CO2eq emissions across the regions are additive. Thus, we have

$$K_m = \sum_i K_i, \tag{35}$$

$$L_m = \sum_i L_i, \tag{36}$$

$$Y_m = \sum_i Y_i, \quad \text{where} \quad Y_i := A_i K_i^\gamma L_i^{1-\gamma}, \tag{37}$$

$$A_m = \frac{Y_m}{K_m^\gamma L_m^{1-\gamma}}, \tag{38}$$

$$\sigma_m = \frac{\sum_i \sigma_i Y_i}{Y_m}. \tag{39}$$

Note that the production function is not scale-invariant:

$$Y(t) = (A(t)K(t))^\gamma (A(t)L_t)^{1-\gamma} \tag{40}$$

$$c \cdot Y(t) = (c \cdot A(t)K(t))^\gamma (c \cdot A(t)L_t)^{1-\gamma} \neq (c \cdot A(t))(c \cdot K(t))^\gamma (c \cdot L_t)^{1-\gamma}, \quad \forall c > 0. \tag{41}$$

Hence, one cannot get the technology after merging multiple regions by simply adding the individual technology levels. Rather, the combined technology factor is imputed from the combined productions, labor, and capital.

Splitting large regions. To avoid huge economies that dominate the fictitious world, we split large economies into pieces based on predetermined fractions c_i and A_is ranging from 0.75 to 2:

$$\sum_i c_i = 1, \tag{42}$$

$$L_i = c_i L_m, \tag{43}$$

$$Y_i = c_i Y_m, \tag{44}$$

$$K_i = \frac{Y_i}{A_i L_i^{1-\gamma}}, \tag{45}$$

$$\sigma_i = \sigma_m. \tag{46}$$

B Model Calibration

The structural parameters of the RICE-N simulation were calibrated to meet the following objectives:

1. Temperatures match the real data in different versions of RICE-N with 1 region, 27 regions, and 266 regions (all raw regions), under 0% and 100% mitigation. To make the computational cost more affordable, we use the 27-region version for the competition.
2. The optimistic-pessimistic temperature outcomes fit the IPCC projections (2 - 7 deg Celsius increase in 100 years). The pessimistic case uses 0% mitigation and 100% savings. The optimistic case uses 100% mitigation and 0% savings.

The parameters that we estimated and the corresponding estimation methods are listed below:

- The dynamic parameters for total factor productivity A: g_A and δ_A.
- The capital K: for the regions whose capital data is not available, we use a KNN regressor [23] to estimate it.
- The dynamic parameters for population L: l_g; similarly, for the regions whose convergence population data is not available, we use a KNN regressor to estimate it.
- The initial carbon intensity σ_0: for the regions whose capital data is not available, we use a KNN regressor to estimate it.
- KNN regressor: Because all regions have GDP and population data, we use them as features. For each region that lacks emission data and capital data, we find the nearest 5 neighbors according to its GDP and population. We use the average of the 5 neighbors' emission data and capital data as the estimated values.

B.1 Population dynamic caliberation

Denoting $L_{\infty;i} := \lim_{t\to\infty} L_{i,t}$, in the limit $t \to \infty$ we have:

$$L_{\infty;i} = L_{\infty;i}\left(\frac{1+L_{a,i}}{1+L_{\infty;i}}\right)^{l_{g;i}}, \tag{47}$$

$$1 = \left(\frac{1+L_{a;i}}{1+L_{\infty}}\right)^{l_{g;i}}. \tag{48}$$

As long as $l_{g;i}$ is not zero, $L_{\infty;i} = L_{a;i}$. Thus, $L_{a;i}$ is the long-term population size and a free parameter that is fitted to data. Assuming $\{L_{i,t}\}_{t=1,2,\ldots}$ is monotonically increasing or decreasing, the absolute value of $l_{g;i}$ represents how fast it converges to $L_{a;i}$. The closer $L_{i,t}$ is to monotonically increasing or monotonically decreasing in the real data, the easier it is to fit $l_{g;i}$ and $L_{a;i}$.

To fit the population parameters, we take logs on both sides of Equation 5:

$$\log L_{i,t+1} - \log L_{i,t} = l_{g;i}(\log(1+L_{a;i}) - \log(1+L_{i,t+1})), \tag{49}$$

where $\log L_{i,t+1} - \log L_{i,t}$ and $\log(1+L_{i,t})$ are given by the data. $\log(1+L_{i,t})$ and $l_{g;i}$ can then be estimated by linear regression.

B.2 Technology dynamic caliberation

We estimate both g_A and δ_A from the existing data $\{A_t\}_{i=1\cdots n}$ by solving a regression problem:

$$g^*_{a;i}, \delta^*_{a;i} = \underset{g_{a;i},\delta_{A,i}}{\arg\max} \, ||A_{i,t} - (\exp\eta + g_{A,i}\exp(-\delta_{A,i}\Delta(t-1)))A_{i,t}||^2. \tag{50}$$

This can be solved by numerical optimization algorithms, e.g., as provided in SciPy [24].

Because the emissions data from the World Bank API do not fit the form of the σ dynamic as assumed by DICE2016, use the DICE2016 parameter values for g_σ and δ_σ.

C WarpDrive Support

Performing multi-agent RL is significantly faster using GPU acceleration. We provide a GPU version of the simulation and an end-to-end GPU-based RL training loop using WarpDrive. For details, see the instruction on the Github repo `https://github.com/mila-iqia/climate-cooperation-competition`.

Table 5: Parameters for 27 regions

Region ID	A_0	K_0	L_0	L_a	$delta_A$	g_A	l_g	$sigma_0$
1	1.872	0.239	476.878	669.594	0.139	0.122	0.034	0.456
2	8.405	3.304	68.395	93.497	0.188	0.103	0.058	0.529
3	3.558	0.109	64.122	135.074	0.161	0.127	0.026	0.816
4	1.927	1.424	284.699	465.308	0.244	0.134	0.024	1.221
5	8.111	0.268	28.141	23.574	0.163	0.106	-0.057	0.290
6	4.217	3.184	548.754	560.054	0.170	0.095	0.080	0.302
7	2.491	0.044	46.489	59.988	0.058	0.049	0.037	0.420
8	2.525	1.080	69.194	100.016	0.346	0.079	0.029	1.010
9	2.460	0.184	513.737	1867.771	1.839	0.462	0.017	0.310
10	12.158	2.642	38.101	56.990	0.131	0.063	0.020	0.350
11	0.993	0.160	522.482	1830.325	0.086	0.065	0.019	0.235
12	5.000	2.289	165.293	230.191	0.183	0.071	0.027	0.419
13	29.854	2.020	165.751	216.927	0.088	0.075	-0.002	0.254
14	23.315	3.039	109.395	143.172	0.088	0.075	-0.002	0.254
15	29.854	0.687	56.355	73.755	0.088	0.075	-0.002	0.254
16	10.922	0.606	705.465	532.497	0.096	0.168	-0.016	0.781
17	9.634	0.608	465.607	351.448	0.096	0.168	-0.016	0.781
18	8.621	0.453	239.858	181.049	0.096	0.168	-0.016	0.781
19	3.190	0.129	690.002	723.513	0.054	0.068	-0.013	0.949
20	2.034	0.381	455.401	477.518	0.054	0.068	-0.013	0.949
21	13.220	16.295	502.410	445.861	0.252	0.074	-0.033	0.170
22	3.190	0.044	234.601	245.994	0.054	0.068	-0.013	0.949
23	6.387	1.094	317.880	287.533	0.194	0.237	-0.053	0.840
24	2.481	0.090	94.484	102.997	0.203	0.201	0.037	1.665
25	10.853	17.554	222.891	168.351	0.005	0.000	-0.012	0.285
26	4.135	1.002	103.294	87.418	0.158	0.123	-0.063	0.601
27	2.716	1.034	573.818	681.210	0.097	0.101	0.043	0.638